LES COURS DE SERGE LANG

Structures algébriques

Mathématiques


InterEditions, Paris

S'il arrive qu'un grand mathématicien ait des préoccupations pédagogiques, il est rare qu'il trouve le temps d'écrire une série d'ouvrages couvrant la quasi-totalité de la formation mathématique de base, c'est-à-dire prenant l'étudiant à la sortie du lycée pour le conduire au seuil de la recherche. C'est pourtant ce qu'a fait Serge Lang.

Contenu de *Structures algébriques*

Les entiers — Groupes — Anneaux — Polynômes — Espaces vectoriels et modules — Théorie des corps (extensions algébriques, plongements, corps de rupture, théorème de Galois, extensions quadratiques et cubiques, résolubilité par radicaux) — Nombres réels et nombres complexes — Ensembles (cardinaux, ensembles bien ordonnés, lemme de Zorn) — Appendices : les entiers naturels, les entiers relatifs, les ensembles infinis.

Structures algébriques est l'ouvrage de l'essentiel : les étudiants tant du premier que du second cycle et les enseignants du second degré trouveront là un des livres dont l'exposé est le plus rapide et le plus clair. La théorie des corps est au centre du livre, les notions que l'on ne peut ignorer sont amplement et clairement illustrées, enfin, au dernier chapitre, sont développés les fondements théoriques plus abstraits, non sans que le lecteur ait auparavant vu l'usage de notions telle que le lemme de Zorn.

Autres ouvrages de Serge Lang parus en français

Algèbre linéaire 1	Espaces vectoriels, matrices, déterminants
Algèbre linéaire 2	Opérateurs, théorie spectrale, algèbre multi-linéaire



STRUCTURES ALGEBRIQUES

SERGE LANG
Yale University

Texte français de

Jean-Marc Braemer
Denis Richard
Université Claude-Bernard, Lyon I



L'édition originale de cet ouvrage par Serge Lang a été publiée aux Etats-Unis
sous le titre *Algebraic Structures* par Addison-Wesley Publishing Company,
Reading, Massachusetts. © 1967.
Ceci est la seule édition française autorisée.

ABCDEFGHIJ-AL-798765

Préface des Traducteurs

S'il arrive qu'un grand mathématicien ait des préoccupations pédagogiques, il est rare qu'il trouve le temps d'écrire une série d'ouvrages couvrant la quasi-totalité de la formation mathématique de base, c'est-à-dire prenant l'étudiant à la sortie du lycée et le conduisant au seuil de la recherche. C'est pourtant ce qu'a fait Serge Lang.

Ces livres reflètent le souci de lier la compréhension à la pratique et sont conçus dans un style remarquablement adapté au travail personnel: la clarté et la simplicité avec laquelle l'auteur sait aborder chaque sujet, sa volonté déliberée de proposer un grand nombre d'exercices d'application immédiate surprendront le lecteur français. Pour ne pas sombrer dans l'abstraction et la généralité abusives, les fondements théoriques suivent bien souvent l'approche pratique d'une notion nouvelle et c'est d'une grande efficacité. Toutefois, ces qualités ne s'exercent pas aux dépens de la profondeur: mathématicien de grande renommée, Serge Lang s'est servi de toutes les notions qu'il enseigne; il en connaît les fondements et les applications, d'où la richesse et l'intérêt exceptionnel des exercices et des exemples proposés. C'est aussi la volonté de l'auteur de présenter chaque notion de façon suffisamment autonome qui permet la publication en fascicules*, s'intégrant parfaitement dans un projet global.

Tout au long de ses études et quel que soit son niveau, l'étudiant trouvera dans les Cours de Serge Lang l'outil qu'il recherche.

Villeurbanne, le 26 mai 1975

Les traducteurs

* ..., and I have frequently committed the crime of *lèse-Bourbaki* by repeating short arguments or definitions to make certain sections or chapters logically independent of each other. (S. Lang, Préface de *Algebra.)*

Préface

Avec *Algèbre linéaire*, ce livre constitue l'exposé du programme d'algèbre des premiers cycles des enseignements supérieurs et de la licence.

Séparer l'algèbre linéaire des autres structures fondamentales correspond aux tendances actuelles de l'enseignement propédeutique, tendances que j'approuve. J'ai construit le présent ouvrage comme se suffisant à lui-même, mais il vaut probablement mieux que les étudiants abordent d'abord le cours d'algèbre linéaire, *avant* de rencontrer les notions plus abstraites de groupes, anneaux et corps, et le développement de leurs propriétés fondamentales. Certaines parties de ce livre sont également traitées dans *Algèbre linéaire*, puisque dans ce dernier livre, j'ai voulu mettre en évidence certains faits concernant les groupes de matrices et les anneaux d'endomorphismes. Cependant, on insiste ici sur des aspects tout à fait différents.

Un cours d'algèbre est également un bon endroit pour initier les étudiants au langage couramment utilisé en mathématiques concernant ensembles et applications, jusqu'au lemme de Zorn inclus. J'ai inséré dans cet esprit, un chapitre sur les ensembles et les cardinaux qui est plus développé que l'habitude. Une des raisons en est que les propositions ici démontrées ne sont pas faciles à trouver dans la littérature, hormis dans des livres de théorie des ensembles, d'un niveau technique élevé. Ainsi le chapitre VII fournira des matériaux intéressants supplémentaires, si l'on en a le temps. Cette partie du livre, ainsi que l'appendice et la construction des nombres réels et complexes, peut également être considérée comme un cours présentant, brièvement et naïvement, des objets mathématiques fondamentaux.

En mathématiques, un texte d'introduction élémentaire, comme celui-ci, devrait être simple et toujours fournir des exemples concrets en même temps qu'il développe la théorie (ce qui explique que j'utilise les nombres réels et complexes comme exemples avant de les avoir rigoureusement construits). Le désir d'éviter des proportions encyclopédiques, des développements trop spécialisés et de faire un ouvrage court expliquent l'omission de quelques théorèmes qui manqueront à certains enseignants souhaitant les inclure dans l'exposé. Les étudiants exceptionnellement doués peuvent toujours suivre des enseignements d'un niveau plus élevé, et on peut utiliser avec eux des textes plus exhaustifs et d'un niveau supérieur faciles à trouver.

SERGE LANG

Sommaire

Chapitre I

Les entiers

Chapitre II

Groupes

Chapitre III

Anneaux

Chapitre IV

Polynômes

Chapitre V

Espaces vectoriels et modules

Chapitre VI

Théorie des corps

Chapitre VII

Les nombres réels et les nombres complexes

Chapitre VIII

Ensembles

Appendice

CHAPITRE I

Les entiers

§1. *Terminologie ensembliste*

Une collection d'objets est appelée un *ensemble*. Un objet de cette collection est aussi appelé un *élément* de l'ensemble. Il est en pratique utile d'utiliser des symboles courts pour désigner certains ensembles. Nous notons, par exemple, $\mathbf{Z}$ l'ensemble de tous les entiers, i.e. tous les nombres du type $0, \pm 1, \pm 2, \ldots$ Au lieu de dire que x est élément de l'ensemble E, nous disons fréquemment que *x est dans E* et nous écrivons $x \in E$. Nous avons par exemple $1 \in \mathbf{Z}$ et aussi $-4 \in \mathbf{Z}$.

Si E et E' sont des ensembles, et si tout élément de E' est un élément de E, nous dirons que E' est un *sous-ensemble* de E. Ainsi l'ensemble des entiers positifs $\{1, 2, 3, \ldots\}$ est un sous-ensemble de l'ensemble de tous les entiers. Dire que E' est un sous-ensemble de E, c'est dire que E' est une partie de E. Remarquons que notre définition du sous-ensemble n'exclut pas la possibilité que $E' = E$. Si E' est un sous-ensemble de E, mais si $E' \neq E$, nous dirons que E' est un sous-ensemble *propre* de E. Ainsi $\mathbf{Z}$ est un sous-ensemble de $\mathbf{Z}$, et l'ensemble des entiers positifs est un sous-ensemble propre de $\mathbf{Z}$. Pour noter que E' est un sous-ensemble de E, nous écrivons $E' \subset E$, et nous disons que E' est *contenu* dans E.

Si E_1 et E_2 sont des ensembles, l'*intersection* de E_1 et de E_2, notée $E_1 \cap E_2$, est l'ensemble des éléments qui sont à la fois dans E_1 et dans E_2. Par exemple, si E_1 est l'ensemble des entiers $\geqq 1$, et si E_2 est l'ensemble des entiers $\leqq 1$, alors

$$E_1 \cap E_2 = \{1\}$$

(ensemble constitué du nombre 1).

La *réunion* de E_1 et de E_2, notée $E_1 \cup E_2$ est l'ensemble des éléments qui sont dans E_1 ou dans E_2. Par exemple, si E_1 est l'ensemble des entiers $\leqq 0$ et si E_2 est l'ensemble des entiers $\geqq 0$, alors $E_1 \cup E_2 = \mathbf{Z}$ est l'ensemble de tous les entiers.

Nous voyons que certains ensembles sont constitués d'éléments décrits par certaines propriétés. Si un ensemble n'a pas d'éléments, il est appelé *ensemble vide*. Par exemple, l'ensemble de tous les entiers x tels que $x > 0$ et $x < 0$ est vide, car il n'existe aucun entier qui soit tel.

Si E et E' sont des ensembles, nous notons $E \times E'$ l'ensemble de tous les couples (x, x') où $x \in E$ et $x' \in E'$.

§2. *Propriétés fondamentales*

On connaît si bien les entiers qu'il serait un peu fastidieux d'en donner une axiomatique tout de suite. Nous supposerons donc que le lecteur est au fait des propriétés élémentaires de l'arithmétique (recouvrant addition, multiplication, inégalités) habituellement enseignées dans l'enseignement secondaire. Le lecteur verra plus tard dans ce livre comment on peut axiomatiser de telles règles (voir, par exemple, le chapitre sur les anneaux pour les règles concernant l'addition et la multiplication, et le chapitre concernant l'ordre pour ce qui est des règles concernant les inégalités).

Donnons explicitement une propriété des entiers que nous prendrons comme axiome sur leur ensemble, et que nous appellerons le *bon ordre*.

Tout ensemble non vide d'entiers $\geqq 0$ *possède un plus petit élément.* (Cela signifie: si S est un ensemble non vide d'entiers positifs ou nul, alors il existe un entiers $n \in S$ tel que $n \leqq x$ pour tout $x \in S$.)

En utilisant ce bon ordre, nous allons démontrer une autre propriété des entiers appelée *propriété de récurrence.* Elle se présente sous plusieurs formes.

Récurrence: *première forme. Supposons que, pour tout entier* $n \geqq 1$, *nous nous sommes donnés une assertion* $A(n)$, *et que nous pouvons démontrer les deux propriétés suivantes:*

(1) *l'assertion* $A(1)$ *est vraie,*

(2) *pour tout entier* $n \geqq 1$, *si* $A(n)$ *est vraie,* $A(n + 1)$ *est vraie.*

Alors, pour tout entier $n \geqq 1$, *l'assertion* $A(n)$ *est vraie.*

Démonstration. Soit E l'ensemble de tous les entiers positifs n pour lesquels l'assertion $A(n)$ est fausse. Nous allons démontrer que E est vide, i.e. qu'il n'y a aucun élément dans E. Supposons qu'il y a un élément dans E. Du fait du bon ordre, il existe un plus petit élément n_0 dans E. Par hypothèse $n_0 \neq 1$, et par conséquent $n_0 > 1$. Puisque n_0 est le plus petit élément de E, $n_0 - 1$ n'est pas dans E; en d'autres termes, l'assertion $A(n_0 - 1)$ est vraie. Mais, d'après la propriété (2), $A(n_0)$, est également vraie puisque

$$n_0 = (n_0 - 1) + 1.$$

Il y a là une contradiction, qui prouve ce que nous voulions démontrer.

Exemple. Démontrons que pour tout entier $n \geqq 1$, $A(n)$:

$$1 + 2 + \cdots + n = \frac{n(n + 1)}{2}$$

est vraie. C'est certainement vrai pour $n = 1$, puisque

$$1 = \frac{1(1 + 1)}{2}$$

Supposons que notre équation soit vérifiée pour un entier $n \geqq 1$. Alors

$$\begin{aligned} 1 + \cdots + n + (n + 1) &= \frac{n(n + 1)}{2} + (n + 1) \\ &= \frac{n(n + 1) + 2(n + 1)}{2} \\ &= \frac{(n + 1)(n + 2)}{2}. \end{aligned}$$

Nous avons ainsi démontré les deux propriétés (1) et (2) pour les assertions notées $A(n)$; nous déduisons par récurrence que $A(n)$ est vraie pour tout entier $n \geqq 1$.

Remarque. Dans l'énoncé de la récurrence, on peut remplacer partout 1 par 0 sans avoir à modifier la démonstration.

La seconde forme de la récurrence est une variante de la première.

Récurrence: *deuxième forme. Supposons que pour tout entier $n \geqq 0$, nous nous sommes donnés une assertion $A(n)$, et que nous pouvons démontrer les deux propriétés suivantes:*

(1′) *l'assertion $A(0)$ est vraie,*
(2′) *pour tout entier $n > 0$, si $A(k)$ est vraie pour tout entier k tel que $0 \leqq k < n$, alors $A(n)$ est vraie.*

Alors, pour tout entier $n \geqq 0$, l'assertion $A(n)$ est vraie.

Démonstration. Soit encore E l'ensemble des entiers $\geqq 0$ pour lesquels l'assertion est fausse. Supposons que E ne soit pas vide, et soit n_0 le plus petit élément de E. Par hypothèse (1′) $n_0 \neq 0$, et puisque n_0 est le plus petit élément de E, pour tout entier k tel que $0 \leqq k < n_0$ l'assertion $A(k)$ est vraie. D'après (2′) $A(n_0)$ est vraie, contradiction qui démontre la seconde forme de la récurrence.

Comme exemple de la seconde forme de la récurrence, nous allons démontré une proposition connue sous le nom d'*algorithme d'Euclide.*

Théorème 1. *Soient m et n des entiers $\geqq 0$, et $m > 0$. Il existe alors des entiers q et r, $\geqq 0$, avec $0 \leqq r < m$ tels que*

$$n = qm + r.$$

Les entiers q et r sont déterminés de manière unique par ces conditions.

Démonstration. Démontrons l'existence de ces entiers par récurrence su n. Si $n = 0$, nous posons $q = r = 0$. Supposons $n > 0$. Si $n < m$, nous posons encore $q = 0$ et $r = n$. Si $n \geqq m$, alors $0 \leqq n - m < n$. Par hypothèse de récurrence, on peut trouver des entiers q_1 et r positifs ou nuls tels que

$$n - m = q_1 m + r.$$

Alors

$$n = m + q_1 m + r = (1 + q_1)m + r.$$

On a ainsi démontré l'existence des entiers q et r cherchés.

Quant à l'unicité, supposons que

$$n = q_1 m + r_1, \quad 0 \leqq r_1 < m$$
$$n = q_2 m + r_2, \quad 0 \leqq r_2 < m$$

Si $r_1 \neq r_2$, par exemple si $r_2 > r_1$, nous obtenons par soustraction

$$(q_1 - q_2)m = r_2 - r_1.$$

Mais $r_2 - r_1 < m$ et $r_2 - r_1 > 0$. Cela est impossible puisque $q_1 - q_2$ est un entier, et qu'ainsi $(q_1 - q_2)m > 0$ implique $(q_1 - q_2)m \geqq m$. Nous en déduisons que $r_1 = r_2$. Mais alors $q_1 m = q_2 m$, donc $q_1 = q_2$. Ce qui prouve l'unicité, et achève la démonstration de notre théorème.

Remarque. Le théorème 1 n'exprime rien d'autre que le résultat de la division de n par m. Nous appellerons r le *reste* de la division de n par m.

EXERCICES

1. Si n et m sont des entiers $\geqq 1$ et si $n \geqq m$, on définit les coefficients binomiaux

$$\binom{n}{m} \quad \frac{n!}{m!(n-m)!}.$$

Comme d'habitude, $n! = n \cdot (n-1) \ldots 1$ est le produit des n premiers nombres entiers. On pose $0! = 1$ et

$$\binom{n}{0} = 1.$$

Démontrez que

$$\binom{n}{m-1} + \binom{n}{m} = \binom{n+1}{m}.$$

2. Démontrez par récurrence que pour tous entiers x et y nous avons

$$(x + y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

Le signe de sommation signifie ici

$$y^n + \binom{n}{1} x y^{n-1} + \binom{n}{2} x^2 y^{n-2} + \cdots + x^n.$$

§3. *Plus grand commun diviseur*

Soient n et d des entiers non nuls. Nous disons que *d divise n* s'il existe un entier q tel que $n = dq$. Nous écrivons alors $d|n$. Si m et n sont des entiers non nuls, nous entendons par *diviseur commun* à m et n tout entier $d \neq 0$ tel que $d|n$ et $d|m$. Par *plus grand commun diviseur* de m et de n, nous entendons tout entier $d > 0$ qui est un diviseur commun et tel que, si $e \neq 0$ est un diviseur de m et de n, alors $e|d$. Nous voyons qu'un tel plus grand commun diviseur existe toujours. On vérifie immédiatement qu'un plus grand commun diviseur est déterminé de manière unique.

Soit J un sous-ensemble de l'ensemble des entiers. Nous disons que J est un *idéal* s'il possède les propriétés suivantes:

L'entier 0 est dans J. Si m et n sont dans J, $m + n$ est dans J. Si m est dans J et si n est un entier quelconque, alors nm est dans J.

Exemple. Soient $m_1, \dots, m_r$ des entiers. Soit J l'ensemble de tous les entiers qui peuvent s'écrire sous la forme

$$x_1 m_1 + \cdots + x_r m_r$$

où $x_1, \dots, x_r$ sont des entiers. On vérifie alors immédiatement que J est un idéal. En effet, si $y_1, \dots, y_r$ sont des entiers

$$(x_1 m_1 + \cdots + x_r m_r) + (y_1 m_1 + \cdots + y_r m_r) = (x_1 + y_1)m_1 + \cdots + (x_r + y_r)m_r$$

est dans J. Si n est un entier

$$n(x_1 m_1 + \cdots + x_r m_r) = nx_1 m_1 + \cdots + nx_r m_r$$

est dans J. Enfin, $0 = 0m_1 + \cdots + 0m_r$ est dans J, et J est un idéal. Nous disons que J est *engendré* par $m_1, \dots, m_r$, et que $m_1, \dots, m_r$, en sont des *générateurs*.

Remarquons que $\{0\}$ est lui-même un idéal, appelé *idéal nul* ou *idéal zéro*. $\mathbf{Z}$ est lui aussi un idéal, dit *idéal unité*.

Théorème 2. *Soit J un idéal de $\mathbf{Z}$. Il existe alors un entier d qui est générateur de J.*

Démonstration. Si J est l'idéal nul, 0 en est un générateur. Supposons $J \neq \{0\}$. Si $n \in J$, $-n = (-1)n$ est aussi dans J, donc J contient au moins un entier positif. Soit d le plus petit entier positif contenu dans J. Nous affirmons que d est un générateur de J. Pour le voir, considérons un $n \in J$ et écrivons $n = dq + r$ avec $0 \leqq r < d$. Alors $r = n - dq$ est dans J, et puisque $r < d$, il s'ensuit $r = 0$. Cela prouve que $n = dq$ et par conséquent que d est un générateur, comme on voulait le démontrer.

Théorème 3. *Soient m_1 et m_2 des entiers positifs. Soit d un générateur positif de l'idéal engendré par m_1 et m_2. Alors, d est un plus grand commun diviseur de m_1 et de m_2.*

Démonstration. Puisque m_1 est dans l'idéal engendré par m_1 et m_2 (car $m_1 = 1m_1 + 0m_2$), il existe un entier q_1 tel que

$$m_1 = q_1 d$$

et par suite d divise m_1. De la même manière, d divise m_2. Soit e un entier non nul divisant à la fois m_1 et m_2 de sorte que, par exemple,

$$m_1 = h_1 e \quad \text{et} \quad m_2 = h_2 e$$

où h_1 et h_2 sont des entiers. Comme d est dans l'idéal engendré par m_1 et m_2, il existe des entiers s_1 et s_2 tels que $d = s_1 m_1 + s_2 m_2$ et par suite

$$d = s_1 h_1 e + s_2 h_2 e = (s_1 h_1 + s_2 h_2) e.$$

Il en résulte que e divise d, et notre théorème est démontré.

Remarque. On peut faire exactement la même démonstration pour plus de deux entiers. Par exemple, si $m_1, \ldots, m_r$ sont des entiers non nuls et si d est un générateur positif de l'idéal engendré par $m_1, \ldots, m_r$, d est un plus grand commun diviseur de $m_1, \ldots, m_r$.

Lorsque le plus grand commun diviseur des entiers $m_1, \ldots, m_r$ est 1, on dit que ces entiers sont *étrangers* ou *premiers entre eux*. Si tel est le cas, il existe des entiers $x_1, \ldots, x_r$ tels que

$$x_1 m_1 + \cdots + x_r m_r = 1$$

puisque 1 appartient à l'idéal engendré par $m_1, \ldots, m_r$.

§4. *Unicité de la décomposition*

Par définition un *nombre premier* p est un entier $\geqq 2$ tel que, si $p = mn$ où m et n sont des entiers positifs, alors $m = 1$ ou $n = 1$. Les plus petits entiers premiers sont 2, 3, 5, 7, 11 ...

Théorème 4. *Tout entier positif $n \geqq 2$ peut s'exprimer comme produit de nombres premiers (non nécessairement distincts)*

$$n = p_1 \cdots p_r,$$

déterminés de manière unique à une permutation près.

Démonstration. Supposons qu'il existe au moins un entier $\geqq 2$ qui ne peut pas s'exprimer comme produit de nombres premiers. Soit m le plus petit élément de l'ensemble des entiers de ce type. En particulier, m n'est pas premier, et nous pouvons écrire $m = de$ pour des entiers d et $e > 1$. Mais alors d et e sont plus petits que m, et étant donnée la manière dont m a été choisi, nous pouvons écrire

$$d = p_1 \cdots p_r \quad \text{et} \quad e = q_1 \cdots q_s$$

où $p_1, \ldots, p_r, q_1, \ldots, q_s$ sont des nombres premiers. Ainsi

$$m = de = p_1 \cdots p_r q_1 \cdots q_s$$

s'exprime comme produit de nombres premiers, contradiction qui démontre que tout nombre entier positif $\geqq 2$ peut s'exprimer comme produit de nombres premiers.

Il faut maintenant montrer l'unicité de cette décomposition, ce qui nécessite le lemme suivant:

Lemme. *Soit p un nombre premier et soit m et n des entiers non nuls tels que p divise mn. Alors $p|m$ ou $p|n$.*

Démonstration. Supposons que p ne divise pas m. Le plus grand commun diviseur de p et de m est alors 1, et il existe des entiers a et b tels que

$$1 = ap + bm.$$

(Utilisez le théorème 3.) En multipliant par n, il vient

$$n = nap + bmn.$$

Mais il existe un entier c tel que $mn = pc$ et par suite

$$n = (na + bc)p,$$

p divise donc n, ce qu'il fallait démontrer.

On va appliquer ce lemme lorsque p divise un produit de nombres premiers $q_1 \cdots q_s$. Dans ce cas, p divise q_1 ou p divise $q_2 \cdots q_s$. Si p divise q_1, alors $p = q_1$. Sinon, on procède par récurrence pour en conclure que dans tous les cas, il existe un entier i tel que $p = q_i$.

Supposons maintenant que nous avons deux produits de nombres premiers

$$p_1 \cdots p_r = q_1 \cdots q_s.$$

D'après ce que nous venons de voir, nous pouvons ré-indexer les entiers $q_1, \ldots, q_s$ et supposer que $p_1 = q_1$. En simplifiant par q_1, on obtient

$$p_2 \cdots p_r = q_2 \cdots q_s.$$

En procédant par récurrence, nous en concluons, par ré-indexation des entiers $q_1, \ldots, q_s$, que nous avons $r = s$, et $p_i = q_i$ pour tout i. Cela prouve l'unicité.

Lorsqu'on exprime des entiers comme produits de nombres premiers, il convient de regrouper les facteurs premiers égaux. Soit donc n un entier > 1, et soient $p_1, \ldots, p_r$ les nombres premiers *distincts* divisant n. Il existe alors des entiers uniques $m_1, \ldots, m_r > 0$ tels que

$$n = p_1^{m_1} \cdots p_r^{m_r}.$$

Nous faisons la convention usuelle que pour tout entier non nul x, $x^0 = 1$. Ainsi, étant donné un entier positif n, nous pouvons écrire n comme produit de puissances de nombres premiers distincts $p_1, \ldots, p_r$:

$$n = p_1^{m_1} \cdots p_r^{m_r},$$

où les exposants $m_1, \ldots, m_r$ sont des entiers positifs ou nuls, et déterminés de manière unique.

L'ensemble des fractions m/n où $n \neq 0$ est appelé ensemble des *nombres rationnels* et est noté $\mathbf{Q}$. Nous supposons pour l'instant que le lecteur connaît cet ensemble. Nous montrerons plus loin comment construire $\mathbf{Q}$ à partir de $\mathbf{Z}$, et comment en établir les propriétés.

Soit $a = m/n$ un nombre rationnel, $n \neq 0$, et supposons $a \neq 0$, donc $m \neq 0$. Soit d le plus grand commun diviseur de m et de n. Nous pouvons donc écrire $m = dm'$ et $n = dn'$, où m' et n' sont premiers entre eux. Ainsi

$$a = \frac{m'}{n'}.$$

Si nous exprimons maintenant $m' = p_1^{i_1} \cdots p_r^{i_r}$ et $n' = q_1^{j_1} \cdots q_s^{j_s}$ comme produits de nombres premiers, nous obtenons une factorisation de a lui-même, et nous remarquons qu'aucun p_ν n'est égal à un q_μ.

Si un nombre rationnel est exprimé sous la forme m/n où m et n sont des entiers, $n \neq 0$ et m et n premiers entre eux, nous disons que n est le *dénominateur* du nombre rationnel et que n en est le *numérateur.* À l'occasion et par abus de langage, lorsqu'on écrit un quotient m/n où m et n ne sont pas nécessairement premiers entre eux, on appelle encore n un dénominateur de la fraction.

§5. *Relations d'équivalence et congruences*

Soit E un ensemble. Nous entendons par *relation d'équivalence* dans E une relation notée $x \sim y$, entre certaines paires d'éléments de E, satisfaisant aux conditions suivantes:

RE 1. *On a* $x \sim x$ *pour tout* $x \in E$.
RE 2. *Si* $x \sim y$ *et* $y \sim z$, *alors* $x \sim z$.
RE 3. *Si* $x \sim y$, *alors* $y \sim x$.

Supposons que nous avons une telle relation d'équivalence dans E. Étant donné un élément x de E, soit C_x l'ensemble de tous les éléments de E équivalents à x. Tous les éléments de C_x sont alors équivalents entre eux, comme on le voit à partir des trois propriétés précédentes. (Vérifiez-le en détail.) De plus, vous vérifiez aussi immédiatement que si x et y sont des éléments de E, alors ou bien $C_x = C_y$, ou bien C_x et C_y n'ont aucun élément commun. Chaque C_x est appelé *classe d'équivalence.* Nous voyons que notre relation d'équivalence détermine une décomposition de E en classes d'équivalence disjointes. Tout élément d'une classe est appelé *représentant* de cette classe.

Le premier exemple que nous donnons de cette notion de relation d'équivalence est la notion de congruence. Soit n un entier. Soient x et y des entiers. Nous disons que *x est congru à y modulo n* s'il existe un entier m tel que $x - y = mn$. Cela signifie que $x - y$ se trouve dans l'idéal engendré par n. Si $n \neq 0$, cela signifie aussi que $x - y$ est divisible par n. Nous notons la relation de congruence de la manière suivante:

$$x \equiv y \pmod{n}.$$

On vérifie immédiatement que c'est une relation d'équivalence, c'est-à-dire que les propriétés suivantes sont vérifiées:

(a) On a $x \equiv x \pmod{n}$.
(b) Si $x \equiv y$ et $y \equiv z \pmod{n}$, alors $x \equiv z \pmod{n}$.
(c) Si $x \equiv y \pmod{n}$, alors $y \equiv x \pmod{n}$.

Les congruences satisfont en outre aux propriétés:

(d) Si $x \equiv y \pmod{n}$ et si z est un entier, alors $xz \equiv yz \pmod{n}$.
(e) Si $x \equiv y \pmod{n}$ et si $x' \equiv y' \pmod{n}$, alors $xx' \equiv yy' \pmod{n}$. De plus $x + x' \equiv y + y' \pmod{n}$.

Nous allons donner une démonstration de la première partie de (e) à titre d'exemple. On peut écrire

$$x = y + mn \quad \text{et} \quad x' = y' + m'n$$

pour des entiers m et m'. Alors

$$xx' = (y + mn)(y' + m'n) = yy' + mny' + ym'n + mm'nn,$$

et l'on voit inmédiatement que l'expression de droite est égale à

$$yy' + wn$$

où w est un entier, de sorte que $xx' \equiv yy' \pmod{n}$, comme on le voulait.

Définissons les entiers *pairs* comme ceux congrus à 0 mod 2. Ainsi n est pair si, et seulement, s'il existe un entier m tel que $n = 2m$. Définissons les entiers *impairs* comme ceux qui ne sont pas pairs. On montre trivialement qu'un entier impair n peut s'écrire sous la forme $2m + 1$, où m est un entier.

Exercices

1. Soient n et d des entiers positifs et supposons que $1 < d < n$. Montrez que n peut s'écrire sous la forme

$$n = c_0 + c_1 d + \cdots + c_k d^k$$

où les c_i sont des entiers tels que $0 \leqq c_i < d$, déterminés de manière unique. [*Indication:* pour démontrer l'existence, écrire $m = qd + c_0$ par division euclidienne, puis procéder par récurrence. Pour démontrer l'unicité, procéder par récurrence en supposant que $c_0, \ldots, c_r$ sont déterminés de manière unique pour prouver que c_{r+1} l'est aussi.]

2. Soient m et n des entiers non nuls écrits sous la forme

$$m = p_1^{i_1}, \ldots, p_r^{i_r} \qquad \text{et} \qquad n = p_1^{j_1}, \ldots, p_r^{j_r},$$

où i_ν et j_μ sont des entiers positifs ou nuls, et $p_1, \ldots, p_r$ des nombres premiers distincts.

(a) Montrez que le p.g.c.d. de m et de n peut s'exprimer comme un produit $p_1^k 1, \ldots, p_r^k r$ où $k_1, \ldots, k_r$ sont des entiers $\geqq 0$. Exprimez k_ν en fonction de i_ν et de j_ν.

(b) Définissez la notion de plus petit commun multiple, et exprimez le plus petit commun multiple de m et de n comme un produit $p_1^{k_1} \cdots p_r^{k_r}$ où les k_ν sont des entiers $\geqq 0$. Exprimez k_ν en fonction de i_ν et de j_ν.

3. Donnez le p.g.c.d. et le p.p.c.m. des couples d'entiers suivants: (a) $5^3 2^6 3$ et 225; (b) 248 et 28.

4. Montrez qu'il existe une infinité de nombres premiers. [*Indication:* étant donné un nombre premier P, soit $N = 2 \cdot 3 \cdot 5 \cdots P + 1$ un poduit contenant comme facteur tous les nombres premiers $\leqq P$. Montrez que tout nombre premier divisant n est plus grand que P.]

5. Soit n un entier $\geqq 2$.

(a) Montrez que tout entier x est congru mod n à un unique entier m tel que $0 \leqq m < n$.

(*b*) Montrez que tout entier $x \neq 0$, premier avec n, est congru à un unique entier m, premier avec n, tel que $0 < m < n$.

(c) Soit $\varphi(n)$ le nombre des entiers m, premiers avec n, tels que $0 < m < n$. La fonction φ est appelée fonction d'Euler. Si $n = p$ est un nombre premier, qu'est-ce-que $\varphi(p)$?

(d) Calculez $\varphi(n)$ pour tout entier n tel que $1 \leqq n \leqq 10$.

6. Soient n et n' deux entiers positifs premiers entre eux. Soient a et b des entiers. Montrez que les congruences

$$x \equiv a \pmod{n}$$
$$x \equiv b \pmod{n'}$$

peuvent être simultanément résolues dans **Z**, en x.

7. Soient a et b des entiers non nuls premiers entre eux; montrez que $1/ab$ peut s'écrire sous la forme

$$\frac{1}{ab} = \frac{x}{a} + \frac{y}{b}$$

où x et y sont des entiers.

8. Montrez que tout nombre rationnel $a \neq 0$ peut s'écrire sous la forme

$$a = \frac{x_1}{p_1^{r_1}} + \cdots + \frac{x_n}{p_n^{r_n}},$$

où $x_1, \ldots, x_n$ sont des entiers, $p_1, \ldots, p_n$ sont des nombres premiers distincts et $r_1, \ldots, r_n$ sont des entiers $\geqq 0$.

9. Soit p un nombre premier et soit n un entier tel que $1 \leqq n \leqq p - 1$. Montrez que le coefficient binomial $\binom{p}{n}$ est divisible par p.

10. Montrez que pour tous les entiers x et y

$$(x + y)^p \equiv x^p + y^p \pmod{p}.$$

CHAPITRE II
Groupes

§1. *Groupes et exemples de groupe*

Un *groupe* G est un ensemble sur lequel on s'est donné une règle (appelée loi de composition) qui permet d'associer à chaque couple d'éléments (x, y) de G un élément de G noté xy, possédant les propriétés suivantes:

GR 1. *Pour tout x, tout y et tout z de G, on a l'associativité, c'est-à-dire*

$$(xy)z = x(yz).$$

GR 2. *Il existe un élément e de G tel que $ex = xe = x$ pour tout x de G. Cet élément est dit élément unité ou élément neutre du groupe.*

GR 3. *Pour tout x de G, il existe un élément y de G tel que $xy = yx = e$.*

A strictement parler, nous dirons que G est un groupe *multiplicatif.* Si nous notons $x + y$ l'élément de G associé au couple (x, y), nous écrivons GR 1 sous la forme

$$(x + y) + z = x + (y + z).$$

GR 2 sous la forme: il existe un élément 0 tel que

$$0 + x = x + 0 = x$$

pour tout x de G, et GR 3 sous la forme: étant donné x de G, il existe un élément y de G tel que

$$x + y = y + x = 0.$$

Avec ces notations, G est appelé groupe *additif.* Nous n'utilisons la notation $+$ que lorsque le groupe satisfait la règle supplémentaire

$$x + y = y + x$$

pour tout x et tout y de G. En notation multiplicative, cela s'écrit $xy = yx$ pour tout x et tout y de G; si G possède cette propriété, nous disons que G est un groupe *commutatif* ou *abélien.*

Nous allons maintenant donner des exemples de groupes. Beaucoup d'entre eux concernent des notions que le lecteur a sans aucun doute déjà rencontrées.

Exemple 1. Soit **Q** l'ensemble des nombres rationnels, i.e. l'ensemble de toutes les fractions m/n, où m et n sont des entiers, et $n \neq 0$. **Q** est alors un groupe pour l'addition. En outre, les éléments non nuls de **Q** forment un groupe pour la multiplication.

Exemple 2. Les nombres réels et les nombres complexes forment des groupes pour l'addition. Les nombres réels non nuls et les nombres complexes non nuls forment des groupes pour la multiplication. Nous notons **R** l'ensemble des nombres réels et **C** celui des nombres complexes.

Exemple 3. Les nombres complexes de module 1 forment un groupe pour la multiplication.

Exemple 4. L'ensemble constitué par les nombres 1 et -1 est un groupe pour la multiplication, et ce groupe possède deux éléments.

Exemple 5. L'ensemble constitué par les nombres $1, i, -1, -i$ est un groupe pour la multiplication. Ce groupe possède quatre éléments.

Exemple 6. *Produit direct de groupes.* Soient G et G' des groupes. Soient $G \times G'$ l'ensemble constitué par tous les couples (x, x') où $x \in G$ et $x' \in G'$. Si (x, x') et (y, y') sont de tels couples, on définit leur produit par $(xy, x'y')$. $G \times G'$ est alors un groupe.

Il suffit simplement de vérifier que toutes les conditions GR 1, GR 2 et GR 3 sont satisfaites, et nous laissons cela au lecteur. Nous appelons $G \times G'$ produit direct de G et de G'.

On peut aussi faire le produit direct d'un nombre fini de groupes. Ainsi si $G_1, \ldots, G_n$ sont des groupes, nous notons

$$\prod_{i=1}^{n} G_i = G_1 \times \cdots \times G_n$$

l'ensemble de tous les n-uples $(x_1, \ldots, x_n)$ où $x_i \in G_i$. Nous définissons la multiplication composante par composante, et nous voyons immédiatement que $G_1 \times \cdots \times G_n$ est un groupe. Si e_i est l'élément neutre de G_i, alors $(e_1, \ldots, e_n)$ est l'élément neutre du produit.

Exemple 7. L'espace euclidien $\mathbf{R}^n$ n'est rien d'autre que le produit

$$\mathbf{R}^n = \mathbf{R} \times \cdots \times \mathbf{R}$$

où **R** est pris n fois. Nous voyons alors que $\mathbf{R}^n$ est un groupe additif.

Un groupe réduit à un seul élément est dit *trivial*. Un groupe quelconque peut avoir une infinité ou seulement un nombre fini d'éléments. Si G n'a qu'un nombre fini d'éléments, on dit que G est un *groupe fini*, et le nombre des éléments de G est appelé son *ordre*. Le groupe de l'exemple 4 est d'ordre 2, et celui de l'exemple 5 est d'ordre 4.

Dans les exemples 1 à 5 les groupes se trouvent être commutatifs. Nous trouverons plus tard des exemples de groupes non commutatifs, lorsque nous étudierons les groupes de permutations.

Soit G un groupe. Soit $x_1, \ldots, x_n$ des éléments de G. On peut alors former leur produit, que nous définissons par récurrence par

$$x_1 \cdots x_n = (x_1 \cdots x_{n-1})x_n.$$

En utilisant l'associativité (GR 1), on peut montrer qu'on obtient la même valeur quelle que soit la place des parenthèses dans ce produit. Par exemple pour $n = 4$,

$$(x_1x_2)(x_3x_4) = x_1\big(x_2(x_3x_4)\big)$$

et aussi

$$(x_1x_2)(x_3x_4) = \big((x_1x_2)x_3\big)x_4.$$

Nous ne donnons pas la démonstration dans le cas général (qui se fait par récurrence), parce qu'elle entraîne de légères complications de notation, dans lesquelles nous ne voulons pas entrer. Le produit ci-dessus est aussi noté

$$\prod_{i=1}^{n} x_i$$

Lorsque la loi de groupe est notée additivement, nous utilisons le signe somme au lieu du signe produit, de telle sorte que la somme de n termes a l'allure suivante

$$\sum_{i=1}^{n} x_i = (x_1 + \cdots + x_{n-1}) + x_n = x_1 + \cdots + x_n.$$

Le groupe G étant commutatif et sa loi notée additivement, on peut montrer par récurrence que la somme ci-dessus est indépendante de l'ordre dans lequel sont pris $x_1, \ldots, x_n$. Nous en omettons encore la démonstration. Si, par exemple, $n = 4$,

$$\begin{aligned}(x_1 + x_2) + (x_3 + x_4) &= x_1 + (x_2 + x_3 + x_4)\\ &= x_1 + (x_3 + x_2 + x_4)\\ &= x_3 + (x_1 + x_2 + x_4).\end{aligned}$$

Nous allons maintenant démontrer diverses assertions simples concernant les groupes.

Soit G un groupe. *L'élément e de G, dont l'existence est assurée par* GR 2, *est déterminé de manière unique*; *en effet, si e et e' satisfont tous deux à cette condition,*

$$e' = ee' = e.$$

On appelle cet élément l'*élément neutre* de G. On l'appelle *zéro* en notation additive.

Soit $x \in G$. L'élément y tel que $yx = xy = e$ est déterminé de manière unique, car si z satisfait à $zx = xz = e$, on a

$$z = ez = (yz)z = y(xz) = ye = y.$$

Nous appelons y l'*inverse* de x, et le notons x^{-1}. En notation additive, nous écrivons $y = -x$.

Soient G un groupe, et H un sous-ensemble de G. Nous disons que H est un *sous-groupe* s'il contient l'élément neutre, et si, pour tout x et tout y élément de

H, xy et x^{-1} sont aussi des éléments de H. (En notation additive, nous écrivons, $x + y \in H$ et $-x \in H$.) H est alors un groupe pour son propre compte, la loi de composition de H étant la même que celle de G. L'élément neutre de G constitue un sous-groupe, et G est un sous-groupe de lui-même.

Exemple 8. Le groupe additif des nombres rationnels est un sous-groupe du groupe additif des nombres réels. Le groupe des nombres complexes de module 1 est un sous-groupe du groupe multiplicatif du groupe des nombres complexes non nuls. Le groupe $\{1, -1\}$ est un sous-groupe de $\{1, -1, i, -i\}$.

Il existe une méthode générale pour obtenir des sous-groupes d'un groupe. Soit E un sous-ensemble d'un groupe G, ayant au moins un élément. Soit H l'ensemble des éléments de G constitué de tous les produits $x_1 \cdots x_n$ tels que x_i ou x_i^{-1} soit, pour tout i, un élément de E, et contenant aussi l'élément neutre. H est alors de manière évidente un sous-groupe de G, appelé sous-groupe *engendré* par E. Nous disons aussi que E est un ensemble de *générateurs* de H.

Exemple 9. Le nombre 1 est un générateur du groupe additif des entiers. En effet, tout entier peut s'écrire sous la forme

$$1 + 1 + \cdots + 1$$

ou

$$-1 - 1 - \cdots - 1,$$

ou alors c'est l'entier 0.

Remarquons qu'en notation additive, la condition pour que E soit un ensemble de générateurs du groupe est que tout élément non nul du groupe puisse s'écrire

$$x_1 + \cdots + x_n,$$

où soit $x_i \in E$, soit $+x_i \in E$.

Exemple 10. Soit G un groupe, et soit a un élément de G. Si n est un entier positif, on pose

$$a^n = a \cdots a,$$

produit de n termes. On pose $a^0 = e$. L'élément a est alors générateur d'un sous-groupe de G, constitué de tous les éléments $(a^{-1})^n$ et a^m pour tous les entiers m et $n \geqq 0$.

Exercices

1. Soient G un groupe et a, b, c des éléments de G. Montrez que $b = c$ lorsque $ab = ac$.
2. Soient G et G' des groupes finis d'ordre m et n respectivement. Quel est l'ordre de $G \times G'$?
3. Soient $x_1, \ldots, x_n$ des éléments d'un groupe G. Démontrez (par récurrence) que

$$(x_1 \ldots x_n)^{-1} = x_n^{-1} \ldots x_1^{-1}.$$

Comment cela s'exprime-t-il en notation additive? Pour deux éléments x et y de G, on a $(xy)^{-1} = y^{-1}x^{-1}$. Écrivez aussi cela en notation additive.

4. Soient G un groupe et $x \in G$. Supposons qu'il existe un entier $n \geqq 1$ tel que $x^n = e$. Montrez qu'il existe un entier $m \geqq 1$, tel que $x^{-1} = x^m$.

5. Soient G un groupe fini. Montrez qu'étant donné x élément de G, il existe un entier $n \geqq 1$, tel que $x^n = e$.

6. Soient G un groupe fini et E un ensemble de générateurs de G. Montrez que tout élément de G peut s'écrire sous la forme

$$x_1 \ldots x_n, \qquad \text{où} \qquad x_i \in E.$$

7. Il existe un groupe G d'ordre 4 ayant deux générateurs x et y tels que $x^2 = y^2 = e$ et $xy = yx$. Déterminez tous les sous-groupes de G. Montrez que

$$G = \{e, x, y, xy\}.$$

8. Il existe un groupe G d'ordre 8 ayant deux générateurs x et y tels que $x^4 = y^2 = e$ tels que $xy = yx^3$. Montrez que les éléments

$$x^i y^j$$

où i et j sont des entiers tels que $i = 0, 1, 2, 3$ et $j = 0, 1$, sont des éléments distincts de G et constituent par conséquent l'ensemble des éléments de G. Déterminez tous les sous-groupes de G.

9. Il existe un groupe G d'ordre 8 ayant des générateurs notés i, j et k tels que

$$ij = k, jk = i, ki = j,$$
$$i^2 = j^2 = k^2.$$

Notons m l'élément i^2. Montrez que les éléments $e, i, j, k, m, mi, mj, mk$ sont des éléments distincts de G. Déterminez tous les sous-groupes de G. Ce groupe G est appelé le *groupe quaternionique*. On écrit fréquemment

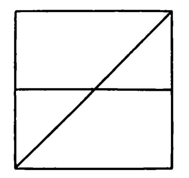
$$-1, -i, -j \text{ et } -k \qquad \text{au lieu de} \qquad m, mi, mj, mk.$$

10. Il existe un groupe G d'ordre 12 ayant des générateurs x et y tels que $x^6 = y^2 = e$ et $xy = yx^5$. Montrez que les éléments

$$x^i y^j$$

où $0 \leqq i \leqq 5$ et $0 \leqq j \leqq 1$, sont des éléments distincts de G. Déterminez tous les sous-groupes de G.

11. Les groupes des exercices 8 à 10 peuvent se représenter comme groupes de symétrie. Par exemple, dans l'exercice 8, soit r la rotation qui envoie chaque sommet du carré



sur le sommet suivant (en prenant par exemple comme sens de rotation le sens inverse de celui des aiguilles d'une montre), et soit s la symétrie par rapport à la diagonale indiquée sur la figure. Démontrez géométriquement que r et s satisfont aux relations de l'exercice 8. Exprimez en fonction des puissances de r et de s, la symétrie par rapport à la droite horizontale indiquée sur la figure.

12. Dans le cas de l'exercice 10, donnez une interprétation géométrique analogue, en prenant un hexagone à la place du carré.

(*Remarque:* les groupes des exercices 11 et 12 peuvent surtout s'interpréter comme groupes de permutations de sommets. (Cf. exercices 12 et 13 du §5.)

13. Soient G un groupe et H un de ses sous-groupes. Soient $x \in G$ et xHx^{-1} le sous-ensemble de G constitué de tous les éléments xyx^{-1} où $y \in H$. Montrez que xHx^{-1} est un sous-groupe de G.

§2. *Applications*

Soient E et E' deux ensembles. Une *application de E dans E'* est un procédé qui, à tout élément de E, associe un élément de E'. Au lieu de dire que f est une application de E dans E', on écrit souvent le symbole $f : E \to E'$.

Si $f : E \to E'$ est une application et si x est un élément de E, on note $f(x)$ l'élément de E' associé à x par f. On appelle $f(x)$ la *valeur* de f en x, ou encore l'*image* de x par f. L'ensemble de tous les éléments $f(x)$ pour tous les x élément de E est appelé l'*image* de f. Si F est une partie de E, l'ensemble des éléments $f(x)$ où $x \in F$, est appelé l'*image* de F et est noté $f(F)$.

Avec les notations précédentes, on écrit souvent $x \mapsto f(x)$ pour noter l'image de x par f. Remarquez que nous distinguons les deux types de flèches $\to$ et $\mapsto$.

Exemple 1. Soient E et E' deux ensembles égaux à $\mathbf{R}$. Soit $f : \mathbf{R} \to \mathbf{R}$ l'application $f(x) = x^2$, i.e. l'application dont la valeur en x est x^2. On peut aussi exprimer cela en disant que f est l'application telle que $x \mapsto x^2$. L'image de f est l'ensemble des nombres réels $\geqq 0$.

Soient $f : E \to E'$ une application et F un sous-ensemble de E. On peut alors définir une application $F \to E'$ par le même procédé $x \mapsto f(x)$, pour $x \in F$. En d'autres termes, on peut considérer f comme définie seulement sur F. Cette application est appelée la *restriction* de f à F et est notée $f \mid F : F \to E'$.

Soient E et E' des ensembles. Une application $f : E \to E'$ est dite *injective*, ou est une *injection*, si, lorsque x et y sont éléments de E, $x \neq y$ implique $f(x) \neq f(y)$.

Exemple 2. L'application f de l'exemple 1 n'est pas injective. On a en effet $f(1) = f(-1)$. Soit $g : \mathbf{R} \to \mathbf{R}$ l'application $x \mapsto x + 1$. L'application g est alors injective, puisque $x \neq y$ implique $x + 1 \neq y + 1$, i.e. $g(x) \neq g(y)$.

Soient E et E' des ensembles. On dit qu'une application $f : E \to E'$ est *surjective*, ou est une *surjection*, si l'*image* $f(E)$ de E est égale à E' tout entier. Cela signifie qu'étant donné un élément quelconque x' de E', il existe un élément x de E tel que $f(x) = x'$. On dit aussi que f est une application de E *sur* E'.

Exemple 3. Soit $f : \mathbf{R} \to \mathbf{R}$ l'application $f(x) = x^2$. L'application n'est alors pas surjective car aucun nombre négatif n'est dans l'image de f.

Soit $g : \mathbf{R} \to \mathbf{R}$ l'application $g(x) = x + 1$. L'application g est surjective car, étant donné un nombre y, on a $y = g(y - 1)$.

Remarque. Soit $\mathbf{R}'$ l'ensemble des nombres réels $\geqq 0$. On peut considérer la correspondance $x \mapsto x^2$ comme une application de $\mathbf{R}$ dans $\mathbf{R}'$. De ce point de

vue l'application est surjective. C'est donc une convention raisonnable que de ne pas identifier cette application à l'application $f : \mathbf{R} \to \mathbf{R}$ définie par le même formule. Pour être parfaitement correct, nous devons inclure dans la notation d'une application l'ensemble de départ et celui d'arrivée et écrire par exemple

$$f^{E}_{E'} : E \to E'$$

au lieu de $f: E \to E'$. En pratique, cette notation est trop lourde, si bien qu'on omet les indices E et E'. Cependant, le lecteur devra garder présent à l'esprit la distinction entre les applications

$$f^{\mathbf{R}}_{\mathbf{R}'}: \mathbf{R} \to \mathbf{R}' \quad \text{et} \quad f^{\mathbf{R}}_{\mathbf{R}}: \mathbf{R} \to \mathbf{R},$$

toutes deux définies par la correspondance $x \mapsto x^2$. La première application est surjective, tandis que la seconde ne l'est pas.

Soit E et E' des ensembles et $f: E \to E'$ une application. On dit que f est *bijective*, ou est une *bijection*, si f est à la fois injective et surjective. Cela signifie qu'étant donné un élément x' de E', il existe un unique élément x de E tel que $f(x) = x'$. (L'élément x existe parce que x est injective, et est unique parce que f est injective.)

Exemple 4. Soit J_n l'ensemble des entiers $\{1, 2, ..., n\}$. Une application bijective $\sigma : J_n \to J_n$ s'appelle une *permutation* des entiers de 1 à n. Ainsi, la permutation σ ci-dessus est en particulier une application $i \mapsto \sigma(i)$. Nous étudions plus en détail les permutations dans la suite de ce chapitre.

Exemple 5. Soit E un ensemble non vide, et soit

$$I : E \to E$$

l'application telle que $I(x) = x$ pour tout $x \in E$. L'application I est appelée *application identique* ou *identité* et est notée *id*. Elle est évidemment bijective.

Soit $f: E \to E'$ une application bijective. On peut définir son *inverse* (ou son *application réciproque*), notée f^{-1}, de la manière suivante:

$$f^{-1}(x') = \text{l'unique élément } x \text{ de } E \text{ tel que } f(x) = x'.$$

Exemple 6. Si $g : \mathbf{R} \to \mathbf{R}$ est l'application telle que $g(x) = x + 1$, alors $g^{-1} : \mathbf{R} \to \mathbf{R}$ est l'application telle que $g^{-1}(x) = x - 1$.

Exemple 7. Soit $\mathbf{R}^+$ l'ensemble des nombres réels positifs (i.e des nombres réels > 0). Soit $h : \mathbf{R}^+ \to \mathbf{R}^+$ l'application $h(x) = x^2$. L'application h est alors bijective et son inverse est l'application racine carrée, i.e. $h^{-1}(x) = \sqrt{x}$ pour tout $x \in \mathbf{R}$, $x > 0$.

Remarque. Si $f: E \to E'$ est une application non nécessairement surjective ou injective, il est souvent commode d'introduire la notion d'*image réciproque* d'un élément de E'. Si donc $y \in E'$, on définit $f^{-1}(y)$ comme l'ensemble de tous les éléments x de E tels que $f(x) = y$. Si y n'est pas dans l'image de f, alors $f^{-1}(y)$ est *vide*. Si y est dans l'image de f, $f^{-1}(y)$ peut contenir plus d'un élément.

Exemple 8. Soit $f : \mathbf{R} \to \mathbf{R}$ l'application $f(x) = x^2$. Alors

$$f^{-1}(1) = \{1, -1\},$$

et $f^{-1}(-2)$ est vide.

Soient A, B et C des ensembles, et soient

$$f : A \to B \quad \text{et} \quad g : B \to C$$

des applications. On peut alors former l'*application composée*

$$g \circ f : A \to C$$

définie par

$$(g \circ f)(x) = g\big(f(x)\big)$$

pour tout x de A.

Exemple 9. Soient $f : \mathbf{R} \to \mathbf{R}$ l'application $f(x) = x^2$ et $g : \mathbf{R} \to \mathbf{R}$ l'application $g(x) = x + 1$. Alors $g\big(f(x)\big) = x^2 + 1$. Remarquons que dans ce cas on peut aussi former $f\big(g(x)\big) = f(x + 1) = (x + 1)^2$, et qu'alors

$$f \circ g \neq g \circ F.$$

La composition des applications est associative. Ce qui signifie que si A, B, C et D sont des ensembles et

$$f : A \to B, \quad g : B \to C \quad \text{et} \quad h : C \to D$$

des applications,

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Démonstration. Elle est très simple. Par définition on a, pour tout élément x de A,

$$\big(h \circ (g \circ f)\big)(x) = h\big((g \circ f)(x)\big) = h\big(g(f(x))\big).$$

D'autre part,

$$\big((h \circ g) \circ f\big)(x) = (h \circ g)\big(f(x)\big) = h\big(g\big(f(x)\big)\big).$$

Par définition, cela signifie que $(h \circ g) \circ f = h \circ (g \circ f)$.

Soient A, B et C des ensembles, $f : A \to B$ et $g : B \to C$ des applications. Si f et g sont injectives, $g \circ f$ est injective. Si f et g sont surjectives, $g \circ f$ est surjective. Si f et g sont bijectives, il en est de même de $g \circ f$.

Démonstration. Supposons d'abord que f et g sont injectives. Soient x et y des éléments de $A, x \neq y$. Alors $f(x) \neq f(y)$ puisque f est injective, donc $g(f(x)) = g\big(f(y)\big)$ puisque g est injective. Par la définition des applications composées, nous concluons que $g \circ f$ est injective. Nous laissons la démonstration de la seconde assertion au lecteur. La troisième est une conséquence des deux premières et de la définition des applications bijectives.

Soient $f: E \to E'$ une application bijective et f^{-1} son inverse. On voit d'après la définition même de l'inverse que

$$f \circ f^{-1} = I_{E'} \quad \text{et} \quad f^{-1}f = I_E,$$

où l'on a indexé l'application i, identiques par leurs ensembles respectifs. En d'autres termes, on a par définition pour tout $x \in E$ et pour tout $x' \in E'$,

$$f(f^{-1}(x')) = x' \quad \text{et} \quad f^{-1}(f(x)) = x.$$

Exemple 10. *Soit E un ensemble non vide et soit G l'ensemble des bijections de E sur lui-même. G est alors un groupe, pour la loi de composition qui est la composition des applications.*

Démonstration. Si $f: E \to E$ et $g: E \to E$ sont deux bijections de E sur lui-même, l'application composée $g \circ f$ est une application de E dans lui-même et est bijective. La condition GR 1 n'est autre que l'associativité de la composition des applications dans le cas présent. L'élément neutre de GR 2 est l'application identique I. Quant à GR 3, ce n'est rien d'autre que l'existence de l'application inverse, si bien que les trois axiomes sont vérifiés.

Remarquons que le groupe G de l'exemple 10 généralise la notion de permutation et, en fait, nous appelons ce groupe G le *groupe des permutations* de l'ensemble E. En pratique, on ne souhaite pas toujours se limiter aux seules permutations des entiers $\{1, ..., n\}$. On souhaite aussi considérer des permutations d'autres ensembles. En matière de notation si σ et τ sont des permutations, on écrit souvent $\sigma\tau$ au lieu de $\sigma \circ \tau$, pour se rapprocher du formalisme utilisé pour les lois de composition des groupes.

Exercices

1. Soit $f: E \to E'$ une application, et supposons qu'il existe une application $g: E' \to E$ telle que

$$g \circ f = I_E \qquad \text{et} \qquad f \circ g = I_{E'},$$

en d'autres termes, supposons que f possède un inverse. Montrez que f est à la fois injective et surjective.

2. Soient $\sigma_1, ..., \sigma_r$ des permutations d'un ensemble E. Montrez que

$$(\sigma_1 ... \sigma_r)^{-1} = \sigma_r^{-1} ... \sigma_1^{-1}$$

3. Soient E un ensemble non vide et G un groupe. Soit $F(E, G)$ l'ensemble des applications de E dans G. Si $f, g \in F(E, G)$, définissons $fg: E \to G$ comme l'application telle que $(fg)(x) = f(x)g(x)$. Montrez que $F(E, G)$ est un groupe. *Si* G est noté additivement, comment écrivez-vous la loi de composition de $F(E, G)$?

4. Donnez l'exemple de deux permutations des entiers $\{1, 2, 3\}$ qui ne commutent pas.

5. Soient E un ensemble, G un groupe, et $f: E \to G$ une application bijective. Pour tout x et tout y de E, on définit le produit

$$xy = f^{-1}(f(x)f(y)).$$

Montrez que cette multiplication définit sur E une structure de groupe.

6. Soient E et F des ensembles, et $f : E \to F$ une application. Soit B une partie de F. Par définition $f^{-1}(B)$ est l'ensemble de tous les éléments x de E tels que $f(x) \in B$. Démontrez que si B et C sont des parties de F

$$f^{-1}(B \cup C) = f^{-1}(B) \cup f^{-1}(C),$$
$$f^{-1}(B \cap C) \subset f^{-1}(B) \cap f^{-1}(C).$$

§3. *Homomorphismes*

Soient G et G' des groupes. Un *homomorphisme*

$$f: G \to G'$$

de G dans G' est une application possédant les propriétés suivantes: pour tout x et tout y éléments de G, on a

$$f(xy) = f(x)f(y)$$

(en notation additive, $f(x + y) = f(x) + f(y)$).

Exemple 1. Soit G un groupe commutatif. L'application $x \mapsto x^{-1}$ de G dans lui-même est un homomorphisme. En notation additive, cette application est $x \mapsto -x$. La vérification en est immédiate.

Exemple 2. L'application

$$z \mapsto |z|$$

est un homomorphisme du groupe multiplicatif des nombres complexes non nuls dans le groupe multiplicatif des nombres complexes non nuls (en fait, dans le groupe multiplicatif des nombres réels positifs).

Exemple 3. L'application

$$x \mapsto e^x$$

est un homomorphisme du groupe additif des nombres réels dans le groupe multiplicatif des nombres réels positifs. Son application inverse, le logarithme, est également un homomorphisme.

Exemple 4. Soit G un groupe. Soit x un élément de G. Si n est un entier positif, on définit x^n comme le produit

$$xx \cdots x$$

de n facteurs égaux. Si $n = 0$, on pose $x^0 = e$. Si $n = -m$, où m est un entier > 0, on a

$$x^{-m} = (x^{-1})^m.$$

C'est une vérification de routine que de constater que

$$x^{m+n} = x^m x^n$$

pour tous entiers m et n. Comme cette vérification est légèrement fastidieuse, nous l'omettons. Mais remarquons qu'en raison de cette propriété, l'application

$$n \mapsto x^n$$

est un homomorphisme du groupe additif **Z** des entiers dans G. Lorsque G est noté additivement, on écrit nx au lieu de x^n.

Pour être bref, nous disons quelquefois: «soit $f: G \to G'$ un homomorphisme de groupes» au lieu de dire: «soient G et G' des groupes, et soit f un homomorphisme de G dans G'».

Soit $f: G \to G'$ un homomorphisme de groupes et soient e et e' les éléments neutres de G et G' respectivement. Alors $f(e) = e'$.

Démonstration. On a $f(e) = f(ee) = f(e)f(e)$. En multipliant par $f(e)^{-1}$, on obtient le résultat souhaité.

Soit $f: G \to G'$ un homomorphisme de groupes. Soit $x \in G$. Alors

$$f(x^{-1}) = f(x)^{-1}.$$

Démonstration. On a

$$e' = f(e) = f(xx^{-1}) = f(x)f(x^{-1}).$$

Soient $f: G \to G'$ et $g: G' \to G''$ deux homomorphismes de groupes. L'application composée $g \circ f$ est alors un homomorphisme de groupes de G dans G''.

Démonstration. On a

$$(g \circ f)(xy) = g(f(xy)) = g(f(x)f(y)) = g(f(x))g(f(y)).$$

Soit $f: G \to G'$ un homomorphisme de groupes. Par définition, le *noyau* de f est constitué de tous les éléments x de G tels que $f(x) = e'$. *On vérifie trivialement que le noyau est un sous-groupe de G.* (Il contient e puisque nous avons montré que $f(e) = e'$; démontrez les autres propriétés est un exercice trivial.)

Soit $f: G \to G'$ un homomorphisme de groupes. Si le noyau de f est réduit à e tout seul, f est alors injective.

Démonstration. Soient x et y deux éléments de G, et supposons que $f(x) = f(y)$, alors

$$e' = f(x)f(y)^{-1} = f(x)f(y^{-1}) = f(xy^{-1}),$$

Par conséquent $xy^{-1} = e$, donc $x = y$, ce qui montre que f est injective.

Soit $f: G \to G'$ un homomorphisme de groupes. L'image de f est un sous-groupe de G'.

Démonstration. Si $x' = f(x)$, où $x \in G$, et si $y' = f(y)$, où $y \in G$,

$$x'y' = f(xy) = f(x)f(y)$$

est aussi dans l'image de f. Sont également dans l'image, e' et $x'^{-1} = f(x^{-1})$ et l'image de f est par suite un sous-groupe de G'.

Soit $f: G \to G'$ un homomorphisme de groupes. Nous disons que f est un

isomorphisme (ou plus précisément un isomorphisme de groupes) s'il existe un homomorphisme $g : G' \to G$ tel que $f \circ g$ et $g \circ f$ sont les applications identiques de G et G' respectivement.

Exemple 5. La fonction *exp* est un isomorphisme du groupe additif des nombres réels sur le groupe multiplicatif des nombres réels positifs. Son inverse est *log*.

Exemple 6. Soit G un groupe commutatif. L'application

$$f : x \mapsto x^{-1}$$

est un isomorphisme de G sur lui-même. Qu'est ce que $f \circ f$? qu'est ce que f^{-1}?

Un homomorphisme de groupes $f : G \to G'$ qui est injectif et surjectif est un isomorphisme.

Démonstration. Soit $f^{-1} : G' \to G$ l'application réciproque de f. Il suffit de prouver que f^{-1} est un homomorphisme de groupes. Soient x' et y' des éléments de G', et soient x et y dans C tels que $f(x) = x'$ et $f(y) = y'$. Alors $f(xy) = x'y'$. Par conséquent, par définition,

$$f^{-1}(x'y') = xy = f^{-1}(x')f^{-1}(y').$$

Cela prouve, comme on le voulait, que f^{-1} est un homomorphisme.

Par *automorphisme* d'un groupe, on entend un isomorphisme de ce groupe sur lui-même. L'application de l'exemple 6 est un automorphisme du groupe commutatif G. A quoi cela ressemble-t-il en notation additive? Des exemples d'automorphismes seront donnés dans les exercices (cf. exercices 3, 4 et 5).

Nous allons voir maintenant que tout groupe est isomorphe au groupe des permutations d'un ensemble.

Exemple 7. Soit G un groupe. Pour tout $a \in G$, soit

$$T_a : G \to G$$

l'application telle que $T_a(x) = ax$. On appelle T_a la *translation à gauche* par a. Nous affirmons que T_a est une bijection de G sur lui-même, i.e. une permutation de G. Si $x \neq y$, alors $ax \neq ay$ (multipliez à gauche par a^{-1} pour le voir), donc T_a est injective. Elle est surjective car, étant donné $x \in G$, on a

$$x = T_a(a^{-1}x).$$

L'application inverse de T_a est bien sûr $T_{a^{-1}}$. Ainsi, l'application

$$a \mapsto T_a$$

est une application de G dans le groupe des permutations de l'ensemble G. Nous affirmons que c'est un homomorphisme. En effet, si $a, b \in G$, on a

$$T_{ab}(x) = abx = T_a(T_b(x)),$$

de telle sorte que $T_{ab} = T_a T_b$. En outre, on voit immédiatement que cet homomorphisme est injectif. L'application

$$a \mapsto T_a \quad (a \in G)$$

est donc un isomorphisme de G sur un sous-groupe du groupe de toutes les permutations de G. Naturellement, toute permutation n'est pas une translation, i.e. l'image de l'application n'est pas égale au groupe des permutations de G tout entier.

La terminologie de l'exemple 7 provient de la géométrie euclidienne. Soit $G = \mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$. Représentons G par le plan. Les éléments de G sont appelés les vecteurs en dimension 2. Si $A \in \mathbf{R} \times \mathbf{R}$, la translation

$$T_A : \mathbf{R} \times \mathbf{R} \to \mathbf{R} \times \mathbf{R}$$

telle que $T_A(X) = X + A$ pour tout $X \in \mathbf{R} \times \mathbf{R}$, est représentée par la translation usuelle de vecteur A.

Exemple 8. *Le groupe des homomorphismes.* Soient A et B des *groupes abéliens* notés additivement. Notons $\mathrm{Hom}(A, B)$ l'ensemble des homomorphismes de A dans B. Nous pouvons faire de $\mathrm{Hom}(A, B)$ un groupe de la manière suivante. Si f et g sont des homomorphismes de A dans B, par définition, $f + g : A \to B$ est l'application telle que

$$(f + g)(x) = f(x) + g(x)$$

pour tout x de A. Il est très simple de vérifier que les trois axiomes de groupe sont vérifiés. En fait, si f, g et h sont dans $\mathrm{Hom}(A, B)$, on a, pour tout $x \in A$,

$$\big((f + g) + h\big)(x) = (f + g)(x) + h(x) = f(x) + g(x) + h(x),$$

et

$$\big(f + (g + h)\big)(x) = f(x) + (g + h)(x) = f(x) + g(x) + h(x).$$

Par suite $f + (g + h) = (f + g) + h$. On a un élément neutre, à savoir l'application 0 (appelée application zéro) qui à tout élément de A associe l'élément nul de B. La condition GR 2 est évidemment vérifiée. De plus, l'application $-f$ telle que $(-f)(x) = -f(x)$ a la propriété suivante

$$f + (-f) = 0.$$

Enfin, il faut naturellement remarquer que $f + g$ et $-f$ sont des homomorphismes. En effet, pour $x, y \in A$,

$$\begin{aligned}(f + g)(x + y) &= f(x + y) + g(x + y) = f(x) + f(y) + g(x) + g(y)\\ &= f(x) + g(x) + f(y) + g(y)\\ &= (f + g)(x) + (f + g)(y),\end{aligned}$$

si bien que $f + g$ est un homomorphisme. De même,

$$(-f)(x + y) = -f(x + y) = -\big(f(x) + f(y)\big) = -f(x) - f(y),$$

et $-f$ est donc un homomorphisme. Cela prouve que $\mathrm{Hom}(A, B)$ est un groupe.

Exercices

1. Soit $\mathbf{R}^\times$ le groupe multiplicatif des nombres réels non nuls. Décrivez explicitement le noyau de l'homomorphisme valeur absolue

$$x \to |x|$$

de $\mathbf{R}^\times$ dans lui-même. Quelle est l'image de cet homomorphisme?

2. Soit $\mathbf{C}^\times$ le groupe multiplicatif des nombres complexes non nuls. Quel est le noyau de l'homomorphisme module

$$z \to |z|$$

de $\mathbf{C}^\times$ dans $\mathbf{R}^\times$.

3. Soit G un groupe et soit a un élément de G. Soit

$$\sigma_a : G \to G$$

l'application définie par

$$\sigma_a(x) = axa^{-1}.$$

Montrez que l'ensemble de telles applications σ_a, où $a \in G$, est un groupe.

4. Montrez que l'ensemble des automorphismes d'un groupe G est lui-même un groupe noté Aut(G).

5. Les notations étant celles de l'exercice 3, montrez que la correspondance $a \mapsto \sigma_a$ est un homomorphisme de G dans Aut(G). L'image de cet homomorphisme est appelée le groupe des automorphismes *intérieurs* de G. Un automorphisme intérieur de G est donc un automorphisme égal à un σ_a pour un $a \in G$.

6. Soit G un groupe abélien fini d'ordre n, et soient $a_1, \ldots, a_n$ ses éléments. Montrez que le produit $a_1 \ldots a_n$ est un élément dont le carré est l'élément neutre.

7. (a) Soit G un sous-groupe du groupe des permutations d'un ensemble E. Si x et y sont des éléments de E, on dit par définition que x est équivalent à y, s'il existe $\sigma \in G$ telle que $\sigma x = y$ Montrez que c'est une relation d'équivalence.

(b) Soient $x \in E$ et G_x l'ensemble de tous les $\sigma \in G$ telles que $\sigma x = x$. Montrez que G_x est un sous-groupe de G.

(c) Si $\sigma \in G$ est telle que $\sigma x = y$, quelle relation y a-t-il entre G_x et G_y?

8. Montrez que tout groupe d'ordre 4 est commutatif.

9. Soit G un groupe commutatif et soit n un entier positif. Montrez que l'application $x \mapsto x^n$ est un homomorphisme de G dans lui-même.

§4. *Classes d'équivalence et sous-groupes distingués*

Soient G un groupe, H un sous-groupe de G et a un élément de G. L'ensemble des éléments ax, où $x \in H$, est appelé une *classe* modulo H dans G. On la note aH.

En notation additive, une classe modulo H s'écrit $a + H$.

Puisqu'un groupe G peut ne pas être commutatif, nous disons en fait que aH est une classe *à gauche* modulo H. On peut, de la même façon, définir les classes *à droite* modulo H, mais dans ce qui suit, sauf mention expresse du contraire, *classe* signifie classe à gauche.

Théorème 1. *Soient aH et bH des classes suivant H dans le groupe G. Ces deux classes sont égales ou disjointes.*

Démonstration. Supposons que aH et bH ont un élément commun. Nous allons démontrer qu'elles sont égales. Soient x et y des éléments de H tels que $ax = by$. Alors $a = byx^{-1}$. Mais yx^{-1} est un élément de H. Si ax' est un élément quelconque de aH, où x' appartient à H,

$$ax' = b(yx^{-1})x'.$$

Puisque $(yx^{-1})x'$ est dans H, nous concluons que ax' est dans bH. Il en résulte que aH est contenue dans bH. De la même manière, bH est contenue dans aH, et par conséquent nos classes sont égales.

Théorème 2. *Soient G un groupe et H un sous-groupe fini de G. Le nombre des éléments d'une classe aH est égal au nombre des éléments de H.*

Démonstration. Soient x et x' des éléments distincts de H; ax et ax' sont alors distincts, car si $ax = ax'$, en multipliant à gauche par a^{-1}, on voit que $x = x'$. Si $x_1, \dots, x_n$ sont les éléments distincts de H, $ax_1, \dots, ax_n$ sont les éléments distincts de aH, et notre assertion en résulte.

Soient G un groupe et H un sous-groupe. Le nombre des classes modulo H dans G est appelé l'*indice* de H dans G. Cet indice peut naturellement être infini. Si G est un groupe fini, l'indice de tout sous-groupe est alors fini. L'indice d'un sous-groupe H est noté $(G : H)$.

Corollaire. *Soient G un groupe fini et H un sous-groupe. Alors*

$$\text{ordre de } G = (G : H)(\text{ordre de } H).$$

Démonstration. Tout élément de G appartient à une classe (a appartient précisément à la classe aH puisque $a = ae$). D'après le théorème 1, tout élément exactement a une classe, et, d'après le théorème 2, deux classes quelconques ont le même nombre d'éléments. La formule de notre corollaire est alors évidente.

Le corollaire montre aussi que l'ordre d'un sous-groupe d'un groupe fini divise l'ordre du groupe.

Exemple 1. Soit S_n le groupe des permutations de $\{1, \dots, n\}$. Soit H le sous-ensemble de S_n constitué de toutes les permutations σ telles que $\sigma(n) = n$ (i.e. toutes les permutations laissant n fixe). Il est clair que H est un sous-groupe et nous pouvons considérer H comme le groupe de permutations S_{n-1} (on suppose $n > 1$). Nous voulons décrire toutes les classes modulo H. Pour tout entier i tel que $1 \leqq i \leqq n$, soit τ_i la permutation telle que $\tau_i(n) = i$, $\tau_i(i) = n$, et τ_i laissant fixes tous les entiers autres que n et i. Nous affirmons que les classes

$$\tau_1 H, \dots, \tau_n H$$

sont distinctes, et constituent toutes les classes selon H dans S_n. Pour le voir, soit $\sigma \in S_n$, et supposons que $\sigma(n) = i$. Alors

$$\tau_i^{-1}\sigma(n) = \tau_i^{-1}(i) = n.$$

Par conséquent, $\tau_i^{-1}\sigma$ est dans H et σ dans $\tau_i H$. Nous avons montré que tout élément de G appartient à une classe $\tau_i H$, et donc que $\tau_1 H, \ldots, \tau_n H$ constituent toutes les classes selon H. Il nous faut encore montrer que ces classes sont distinctes. Si $i \neq j$, pour toute permutation $\sigma \in H$, $\tau_i\sigma(n) = \tau_i(n) = i$ et $\tau_j\sigma(n) = \tau_j(n) = j$. Par suite, $\tau_i H$ et $\tau_j H$ ne peuvent avoir d'élément en commun, pyisque des éléments de $\tau_i H$ et *de* $\tau_j H$ ont des effets distincts sur n. Ce qui prouve notre assertion.

Du corollaire du théorème 2, nous déduisons que

$$\text{ordre de } S_n = n \cdot \text{ordre de } S_{n-1}.$$

Par récurrence, nous voyons immédiatement que

$$\text{ordre de } S_n = n! = n(n-1)\cdots 1.$$

Théorème 3. *Soit $f: G \to G'$ un homomorphisme de groupes. Soit H son noyau, et soit a' un élément de G' appartenant à l'image de f, à savoir: $a' = f(a)$ pour un a de G. L'ensemble des éléments x de G tels que $f(x) = a'$ est exactement la classe aH qui est égale à Ha.*

Démonstration. Soit $x \in aH$, de telle sorte que $x = ah$ pour un $h \in H$. alors

$$f(x) = f(a)f(h) = f(a).$$

Réciproquement, supposons que $x \in G$ et que $f(x) = a'$. Alors

$$f(a^{-1}x) = f(a)^{-1}f(x) = a'^{-1}a' = e'.$$

Il en résulte que $a^{-1}x$ est dans le noyau de H, c'est-à-dire que $a^{-1}x = h$ pour un élément h de H. Donc $x = ah$, et ainsi le noyau est égal à aH. De la même façon, on montre que le noyau est égal à Ha.

Soit $f: E \to E'$ une application. Si x' est un élément de E', on note $f^{-1}(x')$ l'ensemble des éléments x de E tels que $f(x) = x'$, et l'on appelle cet ensemble l'*image réciproque* de x par f. Cet ensemble contient en général plus d'un élément. Dans le théorème 3, on peut dire que l'image réciproque d'un élément a' de G' est une classe de G.

Soit G un groupe et H un sous-groupe. Par définition H est un sous-groupe *distingué* de G, si $aH = Ha$ pour tout $a \in G$, ou, de façon équivalente, si $aHa^{-1} = H$ pour tout $a \in G$. Nous avons simplement vu que le noyau d'un homomorphisme est un sous-groupe distingué. Pour prouver la réciproque, nous avons besoin de notations convenables. Soient E et E' des sous-ensembles d'un groupe G. Par définition EE' est l'ensemble des éléments xx' où $x \in E$ et $x' \in E'$.
Il est alors facile de voir que si E_1, E_2, E_3 sont trois sous-ensembles de G, alors

$$(E_1E_2)E_3 = E_1(E_2E_3).$$

Ce produit est tout simplement formé des éléments xyz, où $x \in E_1$, $y \in E_2$ et $z \in E_3$.

Exemple 2. Montrez que si H est un sous-groupe de G, $HH = H$.

***Théorème* 4.** *Soient G un groupe et H un sous-groupe. Si aH et bH sont des classes modulo H, le produit $(aH)(bH)$ est aussi une classe, et l'ensemble des classes est un groupe pour le produit qu'on vient de définir.*

Démonstration. On a $(aH)(bH) = aHbH = abHH = abH$. Le produit de deux classes est donc une classe. La condition GR 1 est satisfaite en vertu des remarques précédentes sur la multiplication des sous-ensembles de G. La condition GR 2 est satisfaite, l'élément neutre étant la classe $eH = H$ lui-même (le vérifier en détail). La condition GR 3 est satisfaite, l'inverse de aH étant $a^{-1}H$ (le vérifier aussi en détail). Le théorème 4 est ainsi démontré.

Le groupe des classes du théorème 4 est appelé le *groupe quotient* de G par H et est noté G/H. Remarquons que c'est le groupe des classes à droite ou à gauche, puisqu'il n'y a pas de différence entre elles, par suite de l'hypothèse faite sur H. Soulignons que c'est cette hypothèse qui nous permet de définir la multiplication des classes. Si la condition $xH = Hx$, pour tout $x \in G$, n'est pas satisfaite, on ne peut pas définir le groupe quotient.

***Corollaire* 1.** *Soient G un groupe et H un sous-groupe distingué. Soit G/H le groupe quotient, et soit*

$$f: G \to G/H$$

l'application qui à tout $a \in G$ associe la classe $f(a) = aH$. Alors f est un homomorphisme dont le noyau est précisément H.

Démonstration. Le fait que f est un homomorphisme n'est qu'une autre façon d'écrire la définition du produit des classes. Quant au noyau de f, il est clair que tout élément de H lui appartient. Réciproquement, si $x \in G$, et si $f(x) = xH$ est l'élément neutre de G/H, c'est la classe H elle-même, si bien que $xH = H$. Cela signifie que $xe = x$ est un élément de H, donc que H est égal au noyau de f, comme on le voulait.

Nous appelons l'homomorphisme f du corollaire 1, l'*homomorphisme canonique* du groupe G sur le groupe quotient G/H.

***Corollaire* 2.** *Soit $f: G \to G'$ un homomorphisme de groupes dont le noyau est H. Le groupe G/H est alors isomorphe à l'image de f par l'application $aH \mapsto f(aH)$, i.e. par l'application qui associe à toute classe aH l'élément $f(aH)$ de G'.*

Démonstration. Toutes les étapes de la démonstration sont essentiellement triviales et nous les laissons en exercices (cf. exercice 11).

Exemple 3. Considérons le sous-groupe $\mathbf{Z}$ du groupe additif des nombres réels $\mathbf{R}$. Le groupe quotient $\mathbf{R}/\mathbf{Z}$ est parfois appelé le *cercle*. Deux éléments x et y de $\mathbf{R}$ sont dits *congrus* modulo $\mathbf{Z}$, si $x - y \in \mathbf{Z}$. Cette congruence est une relation d'équivalence et les classes d'équivalence sont précisément les classes selon $\mathbf{Z}$ dans $\mathbf{R}$. Si $x \equiv y \pmod{\mathbf{Z}}$, alors $e^{2\pi ix} = e^{2\pi iy}$, et réciproquement. L'application

$$x \to e^{2\pi ix}$$

est un isomorphisme de $\mathbf{R}/\mathbf{Z}$ sur le groupe des nombres complexes de module 1. Pour démontrer ces assertions, il faut bien sûr connaître quelques résultats d'analyse concernant la fonction exponentielle.

Exemple 4. Soit $\mathbf{C}^\times$ le groupe multiplicatif des nombres complexes non nuls et $\mathbf{R}^+$ le groupe multiplicatif des nombres réels positifs. Etant donné un nombre complexe α, on peut écrire

$$\alpha = ru$$

où $r \in \mathbf{R}^+$ et u est de module 1. (On a $u = \alpha/|\alpha|$.) Une telle expression de α est déterminée de manière unique, et l'application

$$\alpha \mapsto \frac{\alpha}{|\alpha|}$$

est un homomorphisme de $\mathbf{C}^\times$ sur le groupe des nombres complexes de module 1. Son noyau est $\mathbf{R}^+$, et il en résulte que $\mathbf{C}^\times/\mathbf{R}^+$ est isomorphe au groupe des nombres complexes de module 1 (cf. exercice 11).

Voir les exercices 21 et 22 pour trouver des représentants des exemples 3 et 4.

Exercices

1. Soit $f : G \to G'$ un homomorphisme de noyau H. Supposons G fini; montrez que

$$\text{ordre de } G = (\text{ordre de l'image de } f)(\text{ordre de } H).$$

2. Soit H un sous-groupe de G de telle sorte que $xH \in Hx$ pour tout $a \in G$. Montrez que H est distingué.

3. Soit H un sous-groupe de G, et supposons que $xHx^{-1} = H$, pour tout élément x de G. Alors $x^{-1}Hx = H$, pour tout $x \in G$.

4. Soient G un groupe et H un sous-groupe de G. Montrez que H est distingué si et seulement si $xHx^{-1} \in H$ pour tout $x \in G$.

5. Montrez que si G est commutatif, tout sous-groupe de G est distingué.

6. Soient H_1 et H_2 deux sous-groupes distingués de G. Montrez que $H_1 \cap H_2$ est distingué.

7. Soient $f : G \to G'$ un homomorphisme et H' un sous-groupe de G'. Montrez que $f^{-1}(H')$ est un sous-groupe de G. Montrez que si H' est distingué dans G', $f^{-1}(H')$ est un sous-groupe distingué de G.

8. Soit $f : G \to G'$ un homomorphisme surjectif. Soit H un sous-groupe distingué de G. Montrez que $f(H)$ est un sous-groupe distingué de G'.

9. Soit G un groupe. Par définition, le *centre* de G est l'ensemble de tous les éléments a de G tels que $ax = xa$, pour tout $x \in G$. Montrez que le centre est un sous-groupe et que ce sous-groupe est distingué. Montrez que ce sous-groupe est le noyau de l'homomorphisme de l'exercice 5 du §3.

10. (a) Soient G un groupe commutatif et H un sous-groupe de G. Montrez que G/H est commutatif. (b) Soient G un groupe et H un sous-groupe distingué de G. Montrez que G/H est commutatif si, et seulement si, H contient tous les éléments $xyx^{-1}y^{-1}$, pour x et y éléments de G.

11. Soit $f : G \to G'$ un homomorphisme de groupes, et soit H son noyau. Supposons que G' soit l'image de f; montrez que G/H est isomorphe à G'.

12. Soient G un groupe et H un de ses sous-groupes. Soit N_H l'ensemble de tous les éléments x de G tels que $xHx^{-1} = H$. Montrez que N_H est un groupe contenant H, et que H est distingué dans N_H.

13. Soient G un groupe, H un sous-groupe de G et N un sous-groupe distingué de G. Montrez que NH est un sous-groupe de G et que $NH = HN$.

14. (a) Soit G l'ensemble de toutes les applications de $\mathbf{R}$ dans lui-même de la forme $x \mapsto ax + b$, où $a \in \mathbf{R}$, $a \neq 0$ et $b \in \mathbf{R}$. Montrez que G est un groupe. Notons $\sigma_{a,b}$ une telle application, *i.e.* $\sigma_{a,b}(x) = ax + b$.

(b) A toute application $\sigma_{a,b}$ on associe le nombre a. Montrez que la correspondance

$$\sigma_{a,b} \mapsto a$$

est un homomorphisme de G dans $\mathbf{R}$. Quel est son noyau?

15. Soient G un groupe et H un sous-groupe de G. Si $x, y \in G$, on dit que x est équivalent à y si x appartient à la classe yH. Montrez qu'on définit bien ainsi une relation d'équivalence.

16. Soit G un groupe. Soit E l'ensemble des sous-groupes de G. Si H et K dont des sous-groupes de G, on dit que H est équivalent à K s'il existe un élément x de G tel que $xHx^{-1} = K$. Montrez qu'on définit ainsi une relation d'équivalence sur E.

17. Soient G un groupe et E l'ensemble des sous-groupes de G. Pour tout $x \in G$, soit $f_x : E \to E$ l'application définie par

$$f_x(H) = xHx^{-1}.$$

Montrez que f_x est une permutation de E, et que l'application $x \mapsto f_x$ est un homomorphisme de G dans le groupe des permutations de E.

18. Soient G un groupe et H un sous-groupe de G. Soit E l'ensemble des classes modulo H dans G. Pour tout $x \in G$, soit $g_x : E \to E$ l'application qui à toute classe yH associe la classe xyH. Montrez que g_x est une permutation de E, et que l'application $x \mapsto g_x$ est un homomorphisme de G dans le groupe des permutations de E.

19. Considérons $\mathbf{Z}$ comme un sous-groupe du groupe additif $\mathbf{Q}$ des nombres rationnels. Montrez qu'étant donné un élément $\bar{x}$ de $\mathbf{Q}/\mathbf{Z}$, il existe un entier $n \geqq 1$, tel que $n\bar{x} = 0$.

20. Soit D le sous-groupe de $\mathbf{R}$ engendré par 2π. Soit $\mathbf{R}^+$ le groupe multiplicatif des nombres réels positifs et soit $\mathbf{C}^\times$ le groupe multiplicatif des nombres complexes non nuls. Montrez que $\mathbf{C}^\times$ est isomorphe à $\mathbf{R}^+ \times \mathbf{R}/D$ par l'application

$$(r, \theta) \mapsto re^{i\theta}.$$

(Il faut naturellement utiliser les propriétés de l'exponentielle complexe.)

21. Montrez que toute classe modulo $\mathbf{Z}$ dans $\mathbf{R}$ possède un unique représentant x tel que $0 \leqq x < 1$. [*Indication:* pour tout nombre réel y, soit n l'entier tel que $n \leqq y < n + 1$.]

22. Montrez que toute classe modulo $\mathbf{R}^+$ dans $\mathbf{C}^\times$ possède un unique représentant de module 1.

23. Montrez que tout sous-groupe du groupe des quaternions est distingué (cf. exercice 9 du §1).

24. Déterminez tous les sous-groupes distingués du groupe de l'exercice 8 du §1.

§5. *Groupes de permutations*

Nous allons étudier plus en détail le groupe S_n des permutations de n éléments $\{1, ..., n\} = J_n$, dans ce paragraphe.

Si $\sigma \in S_n$, on rappelle que $\sigma^{-1} : J_n \to J_n$ est la permutation telle que $\sigma^{-1}(k)$ est l'unique entier $j \in J_n$ tel que $\sigma(j) = k$. Une *transposition* τ est une permutation qui échange deux nombres et laisse fixes les autres, i.e. telle qu'il existe des entiers i et j de J_n, $i \neq j$, tels que $\tau(i) = j$, $\tau(j) = i$ et $\tau(k) = k$, pour tout k différent de i et de j. On voit de plus que, si τ est une transposition, $\tau^{-1} = \tau$ et $\tau^2 = I$; en particulier, l'inverse d'une transposition est une transposition.

Théorème 5. *Toute permutation de J_n peut s'exprimer comme produit de transpositions.*

Démonstration. Nous allons démontrer notre assertion par récurrence sur n. Pour $n = 1$, il n'y a rien à démontrer. Soit $n > 1$, et supposons que l'assertion est démontrée pour $n - 1$. Soit σ une permutation de J_n, et soit $\sigma(n) = k$. Soit τ la transposition de J_n telle que $\tau(k) = n$ et $\tau(n) = k$. Alors $\tau\sigma$ est une permutation telle que

$$\tau\sigma(n) = \tau(k) = n.$$

En d'autres termes, $\tau\sigma$ laisse n fixe. On peut donc considérer $\tau\sigma$ comme une permutation de J_{n-1}, et, par hypothèse de récurrence, il existe des transpositions $\tau_1, ..., \tau_s$ de J_{n-1}, laissant n fixe, telles que

$$\tau\sigma = \tau_1 \cdots \tau_s.$$

On peut alors écrire

$$\sigma = \tau^{-1}\tau_1, ..., \tau_s,$$

ce qui démontre notre proposition.

En vue du prochain théorème, il est nécessaire de définir des opérations sur les permutations. Considérons le produit

$$\mathbf{Z} \times \cdots \times \mathbf{Z} = \mathbf{Z}^n,$$

i.e. l'ensemble des n-uples d'entiers. Soit σ une permutation de J_n; σ induit alors une application, notée π_σ, de $\mathbf{Z}^n$ dans lui-même, définie par

$$\pi_\sigma(x_1, ..., x_n) = (x_{\sigma(1)}, ..., x_{\sigma(n)}),$$

pour $x_1, ..., x_n$ éléments de $\mathbf{Z}$. En d'autres termes, π_σ permute simplement les facteurs de $\mathbf{Z}^n$. On vérifie immédiatement que

$$\pi_\sigma \circ \pi_\tau = \pi_{\tau\sigma},$$

pour tout σ et tout τ de S_n. Pour le voir, soit $y_i = x_{\tau(i)}$; alors

$$\begin{aligned}\pi_\sigma \circ \pi_\tau(x_1, \ldots, x_n) &= \pi_\sigma(x_{\tau(1)}, \ldots, x_{\tau(n)})\\ \pi_\sigma \circ \pi_\tau &= \pi_\sigma(y_1, \ldots, y_n)\\ &= (y_{\sigma(1)}, \ldots, y_{\sigma(n)})\\ &= (x_{\tau\sigma(1)}, \ldots, x_{\tau\sigma(n)})\\ &= \pi_{\tau\sigma}(x_1, \ldots, x_n).\end{aligned}$$

Soit $f: \mathbf{Z} \times \cdots \times \mathbf{Z} \to \mathbf{Z}$ une application. On pose

$$f^\sigma = f \circ \pi_\sigma,$$

de telle sorte que

$$f^\sigma(x_1, \ldots, x_n) = f\left(\pi_\sigma(x_1, \ldots, x_n)\right) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

Nous avons alors, par associativité,

$$f^{\sigma\tau} = f \circ \pi_{\sigma\tau} = f \circ \pi_\tau \circ \pi_\sigma = (f^\tau)^\sigma.$$

Si f et g sont deux fonctions de $\mathbf{Z}^n$ dans $\mathbf{Z}$, on peut former leur somme et leur produit de la façon habituelle. La somme $f + g$ est définie par

$$(f + g)(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) + g(x_1, \ldots, x_n)$$

et le produit fg est défini par

$$(fg)(x_1, \ldots, x_n) = f(x_1, \ldots, x_n)g(x_1, \ldots, x_n).$$

Nous affirmons que

$$(f + g)^\sigma = f^\sigma + g^\sigma \quad \text{et} \quad (fg)^\sigma = f^\sigma g^\sigma.$$

Pour le voir, on constate que

$$\begin{aligned}(f + g) \circ \pi_\sigma(x_1, \ldots, x_n) &= (f + g)(x_{\sigma(1)}, \ldots, x_{\sigma(n)})\\ &= f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) + g(x_{\sigma(1)}, \ldots, x_{\sigma(n)})\\ &= f \circ \pi_\sigma(x_1, \ldots, x_n) + g \circ \pi_\sigma(x_1, \ldots, x_n),\end{aligned}$$

ce qui prouve la première affirmation. Pour la seconde, il suffit de remplacer dans les égalités précédentes, les sommes par des produits. Cela prouve ce que nous voulions démontrer. Vérifiez aussi, à titre d'exercice, que

$$(-f)^\sigma = -f^\sigma.$$

Théorème 6. *A toute permutation σ de J_n, on peut associer l'un des deux entiers $+1$ ou -1, noté $\epsilon(\sigma)$, de façon à satisfaire les conditions suivantes:*

(a) *si τ est une transposition $\epsilon(\tau) = -1$,*
(b) *si σ et σ' sont des permutations de J_n, alors*

$$\epsilon(\sigma\sigma') = \epsilon(\sigma)\epsilon(\sigma').$$

Démonstration. Soit Δ la fonction définie par

$$\Delta(x_1, \ldots, x_n) = \prod_{i<j} (x_j - x_i)$$

ce produit étant pris pour tous les couples d'entiers i, j satisfaisant à

$$1 \leqq i < j \leqq n$$

Soit τ une transposition échangeant deux entiers r et s. Supposons par exemple que $r < s$. Nous voulons déterminer

$$\Delta^\tau(x_1, \ldots, x_n) = \prod_{i<j} (x_{\tau(j)} - x_{\tau(i)})$$
$$= \prod_{i<j} (x_j - x_i)^\tau .$$

On a

$$(x_s - x_r)^\tau = (x_r - x_s) = -(x_s - x_r).$$

Si un facteur ne contient ni x_r ni x_s, il reste invariant par application de τ. On peut prendre tous les autres facteurs deux par deux, de l'une des manières suivantes:

$$\begin{array}{ll} (x_k - x_s)(x_k - x_r) & \text{si} \quad k > s, \\ (x_s - x_k)(x_k - x_r) & \text{si} \quad r < k < s, \\ (x_s - x_k)(x_r - x_k) & \text{si} \quad k < r. \end{array}$$

Chacun de ces trois produits reste invariant par application de τ. Nous voyons donc que

$$\Delta^\tau = -\Delta.$$

Soit maintenant σ une permutation quelconque; exprimons σ comme produit de transpositions,

$$\sigma = \bar{\tau}_m \cdots \bar{\tau}_1.$$

On trouve, par récurrence, que

$$\Delta^\sigma = (\Delta^{\tau 1})^{\tau_m \cdots \tau_2} = (-\Delta)^{\tau_m \cdots \tau_2} = (-1)^m \Delta.$$

Ainsi, si nous exprimons σ d'une autre manière comme produit de transpositions,

$$\sigma = \bar{\tau}_1 \cdots \bar{\tau}_k,$$

nous voyons que $(-1)^m \Delta = (-1)^k \Delta$. Il en résulte que $(-1)^m = (-1)^k$, et, par conséquent, que, dans toute décomposition de σ en un produit de transpositions, le nombre des facteurs est toujours soit pair, soit impair. On pose

$$\epsilon(\sigma) = (-1)^m,$$

entier dont nous venons de remarquer qu'il est indépendant de l'expression de σ en produit de transpositions. Si

$$\sigma = \tau_1 \cdots \tau_m \quad \text{et} \quad \sigma = \tau'_1 \cdots \tau'_k,$$

alors

$$\sigma\sigma' = \tau_1 \cdots \tau_m \tau'_1 \cdots \tau'_k,$$

de telle sorte que

$$\epsilon(\sigma\sigma') = (-1)^{m+k} = (-1)^m(-1)^k = \epsilon(\sigma)\epsilon(\sigma').$$

Ce qui prouve notre théorème. Nous avons également montré le

***Corollaire* 1.** *Si une permutation σ de J_n s'exprime comme produit de transpositions:*

$$\sigma = \tau_1 \cdots \tau_s,$$

l'entier s est pair ou impair selon que $\epsilon(\sigma) = 1$ *ou* -1.

***Corollaire* 2.** *Si σ est une permutation de J_n,*

$$\epsilon(\sigma) = \epsilon(\sigma^{-1}).$$

Démonstration. On a

$$1 = \epsilon(id) = \epsilon(\sigma\sigma^{-1}) = \epsilon(\sigma)\epsilon(\sigma^{-1}).$$

Par conséquent, soit $\epsilon(\sigma)$ et $\epsilon(\sigma^{-1})$ sont tous deux égaux à 1, soit ils sont tous deux égaux à -1, ce qu'il fallait démontrer.

En ce qui concerne la terminologie, une permutation est dite *paire* si sa signature est 1; elle est dite *impaire* si sa signature est -1 (la *signature* d'une permutation σ est l'entier $\epsilon(\sigma)$ - N.d.T.). Ainsi, toute transposition est impaire.

Le théorème 6 montre que l'application

$$\epsilon: S_n \to \{1, -1\}$$

est un homomorphisme de S_n sur le groupe constitué des deux éléments 1 et -1. Le noyau de cet homomorphisme est par définition l'ensemble des permutations paires et est appelé le *groupe alterné* A_n. Si τ est une transposition, A_n et τA_n sont évidemment des classes distinctes modulo A_n, et toute permutation appartient soit à A_n, soit à τA_n. (*Démonstration:* si $\sigma \in S_n$ et $\sigma \notin A_n$, alors $\epsilon(\sigma) = -1$; ainsi $\epsilon(\tau\sigma) = 1$ et $\tau\sigma \in A_n$, d'où $\sigma \in \tau^{-1}A_n = \tau A_n$). Par conséquent

$$A_n, \quad \tau A_n$$

sont des classes distinctes modulo A_n dans S_n, et il n'y en a pas d'autres. Puisque A_n est le noyau d'un homomorphisme, A_n est un sous-groupe distingué de S_n. On a $\tau A_n = A_n \tau$, ce qui peut être facilement vérifié directement.

Une permutation σ de $\{1, \ldots, n\}$ se note parfois

$$\begin{bmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{bmatrix}.$$

Ainsi

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}$$

représente la permutation σ telle que $\sigma(1) = 2, \sigma(2) = 1$ et $\sigma(3) = 3$. Cette permutation est, en fait, une transposition

Soient $i_1, \ldots, i_r$ des entiers distincts de J_n. Le symbole

$$(i_1 \cdots i_r)$$

représente la permutation σ telle que

$$\sigma(i_1) = i_2, \quad \sigma(i_2) = i_3, \ldots, \sigma(i_r) = i_1,$$

et qui laisse invariant tous les autres entiers. Par exemple

$$[132]$$

représente la permutation σ telle que $\sigma(1) = 3$, $\sigma(3) = 2$ et $\sigma(2) = 1$, laissant invariant, par ailleurs, tous les autres entiers. Une telle permutation est appelée un *cycle*, et plus précisément, un *r-cycle* (*cycle de longueur r* - N.d.T.).

Si $\sigma = [i_1 \cdots i_r]$, c'est un cycle et on vérifie immédiatement que σ^{-1} est aussi un cycle, et qu'en fait,

$$\sigma^{-1} = [i_r \cdots i_1].$$

Ainsi, si $\sigma = [132]$, $\sigma^{-1} = [231]$.

Remarquons qu'un 2-cycle $[ij]$ n'est rien d'autre qu'une transposition, à savoir la transposition telle que $i \mapsto j$ et $j \mapsto i$.

Un produit de cycles se calcule facilement. Par exemple,

$$[132][34] = [2134].$$

On le voit en revenant à la définition: si $\sigma = [132]$ et $\tau = [34]$, alors, par exemple

$$\begin{aligned} \sigma(\tau(3)) &= \sigma(4) = 4, \\ \sigma(\tau(4)) &= \sigma(3) = 2, \\ \sigma(\tau(2)) &= \sigma(2) = 1, \\ \sigma(\tau(1)) &= \sigma(1) = 3. \end{aligned}$$

Soit G un groupe. On dira que G est *résoluble* s'il existe une suite de sous-groupes

$$G = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_m = \{e\}$$

telle que H_i soit distingué dans H_{i-1}, et telle que le groupe quotient H_{i-1}/H_i soit abélien pour $i = 1, \ldots, m$. Nous allons démontrer que pour $n \geqq 5$, le groupe S_n n'est pas résoluble. Cela nécessite quelques préliminaires.

Théorème 7. *Soit G un groupe et soit H un sous-groupe distingué de G. Alors G/H est abélien si, et seulement si, H contient tous les éléments de la forme $xyx^{-1}y^{-1}$, où $x, y \in G$.*

Démonstration. Soit $f: G \to G/H$ l'homomorphisme canonique. Supposons que G/H soit abélien. Pour tous éléments x et y de G, on a

$$f(xyx^{-1}y^{-1}) = f(x)f(y)f(x)^{-1}f(y)^{-1},$$

et puisque G/H est abélien, l'expression de droite est égale à l'élément neutre de G/H.

Par conséquent, $xyx^{-1}y^{-1}$ appartient à H. Réciproquement, supposons qu'il en soit ainsi pour tous éléments x et y de G. Soient $\bar{x}$ et $\bar{y}$ des éléments de G/H. Puisque f est surjective, il existe des éléments x et y de G tels que $\bar{x} = f(x)$ et $\bar{y} = f(y)$. Soient $\bar{e}$ l'élément neutre de G/H, et e celui de G. Alors

$$\bar{e} = f(e) = f(xyx^{-1}y^{-1}) = f(x)f(y)f(x)^{-1}f(y)^{-1} = \bar{x}\bar{y}\bar{x}^{-1}\bar{y}^{-1}.$$

En multipliant à droite par $\bar{y}$, puis par $\bar{x}$, il vient

$$\bar{y}\bar{x} = \bar{x}\bar{y},$$

et G/H est donc abélien.

***Théorème* 8.** *Si $n \geqq 5$, S_n n'est pas résoluble.*

Démonstration. Nous allons d'abord démontrer que, si H et N sont deux sous-groupes de S_n tels que $N \subset H$ et que N soit distingué dans H, alors si H contient tous les 3-cycles, et si H/N est abélien, N contient tous les 3-cycles. Pour le voir, considérons i, j, k, r et s, cinq entiers distincts compris entre 1 et n, et soit

$$\sigma = [ijk] \quad \text{et} \quad \tau = [krs].$$

Alors

$$\sigma\tau\sigma^{-1}\tau^{-1} = [ijk][krs][kji][srk] = [rki].$$

Puisqu'on a arbitrairement choisi i, j, k, r et s, on voit que les cycles $[rki]$ sont tous dans N, quels que soient les entiers r, k et i distincts, ce qui prouve notre assertion.

Supposons maintenant que nous ayons une chaîne de sous-groupes

$$S_n = H_0 \supset H_1 \supset H_2 \supset \cdots \supset H_m = \{e\}$$

telle que H_i soit distingué dans H_{i-1} et que H_{i-1}/H_i soit abélien, pour $i = 1, \dots, m$. Puisque S_n contient tous les 3-cycles, on déduit que H_1 contient tous les 3-cycles. Par récurrence sur n, on déduit que $H_m = \{e\}$ contient aussi tous les 3-cycles, ce qui est impossible. Une telle chaîne de sous-groupes n'existe donc pas et notre théorème est démontré.

EXERCICES

1. Soient G un groupe, N un sous-groupe distingué et H un sous-groupe de G. Montrez que $H \cap N$ est distingué dans H.

2. Les hypothèses étant celles de l'exercice précédent, montrez que si G/N est abélien, $H/(H \cap N)$ l'est aussi.

3. Soient G un groupe résoluble et H un sous-groupe de G. Montrez que H est résoluble.

4. Soient G un groupe résoluble et $f : G \to G'$ un homomorphisme surjectif. Montrez que G' est résoluble.

5. Déterminez les signatures des permutations suivantes:

(a) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{bmatrix}$

6. Écrivez l'inverse de chacune des permutations de l'exercice 5.

7. Montrez que, pour $n \geqq 2$, le nombre des permutations impaires de $\{1, ..., n\}$ est égale au nombre des permutations paires.

8. Montrez que les groupes S_2, S_3 et S_4 sont résolubles.

9. Soit σ le r-cycle $[i_1 ... i_r]$. Montrez que $\epsilon(\sigma) = (-1)^{r+1}$. [*Indication:* par récurrence.] Si $r = 2, \sigma$ est une transposition. Si $r > 2$,

$$[i_1 ... i_r] = [i_1 i_r][i_1 ... i_{r-1}].$$

10. On dit que deux cycles $[i_1 ... i_r]$ et $[j_1 ... j_s]$ sont *disjoints* si aucun entier i_p n'est égal à un entier j_q. Montrez que toute permutation est égale à un produit de cycles disjoints. [*Indication:* soit σ une permutation. Nous dirons que i est équivalent à j s'il existe un entier $k \geqq 0$, tel que $\sigma^k(i) = j$. Montrez que la relation précédente est d'équivalence et que tout classe d'équivalence détermine un cycle, par exemple $[i\sigma(i)\sigma^2(i)...]$

11. Exprimez les permutations de l'exercice 5 comme produits de cycles disjoints.

12. Montrez que le groupe de l'exercice 8, §1 existe en l'exhibant comme sous-groupe de S_4, de la manière suivante. Soit $\sigma = [1234]$ et $\tau = [24]$; montrez que le groupe engendré par σ et τ est d'ordre 8 et que σ et τ satisfont aux mêmes relations que x et y de l'exercice cité ci-dessus.

13. Montrez que le groupe de l'exercice 10, §1 existe en l'exhibant comme sous-groupe de S_6.

14. Soit n un entier pair et positif. Montrez qu'il existe un groupe d'ordre $2n$ engendré par deux éléments σ et τ tels que $\sigma^n = e = \tau^2$ et $\sigma\tau = \tau\sigma^{n-1}$.

§6. *Groupes cycliques*

L'ensemble **Z** des entiers est un groupe additif. Nous allons déterminer ses sous-groupes. Soit H un sous-groupe de **Z**. Si H n'est pas trivial, soit a le plus petit entier positif appartenant à H. Nous affirmons que H est formé de tous les éléments na, où $n \in \mathbf{Z}$. Pour le montrer, considérons un élément $y > 0$ de H. Il existe des entiers n et r tels $0 \leqq r < a$ et

$$y = na + r$$

Puisque H est un sous-groupe de **Z**, et puisque $r = y - na$, $r \in H$ et par conséquent $r = 0$. Si $y < 0$, on applique le raisonnement précédent à $-y$, qui est dans H, puisque H est un sous-groupe.

Soit G un groupe. Nous disons que G est *cyclique* s'il existe un élément a de G tel que tout élément x de G peut s'écrire sous la forme a^m, où m est un entier (cela revient à dire que l'application $f: \mathbf{Z} \to G$ définie par $f(n) = a^n$ est surjective). Un tel élément a de G s'appelle un *générateur* de G.

Soient G un groupe et $a \in G$. L'ensemble de tous les éléments a^n $(n \in \mathbf{Z})$ est evidemment un sous-groupe cyclique de G. Si m est un entier tel que $a^m = e$ et $m > 0$, nous disons que m est un *exposant* de a.

Soient G un groupe et a un élément de G. Soit $f: \mathbf{Z} \to G$ l'homomorphisme tel que $f(n) = a^n$, et soit H le noyau de f. Deux cas peuvent se produire:

(i) le noyau est trivial. L'application f est alors un isomorphisme de **Z** sur le sous-groupe cyclique de G engendré par a, puisque f est injective et que l'image

de f est précisément égale à ce sous-groupe. De plus, ce sous-groupe est cyclique et infini. Si a engendre G, G est cyclique. Nous disons aussi que l'élément a est de *période infinie*, (ou d'*ordre infini* - N.d.T.).

Exemple 1. Le nombre 2 engendre un sous-groupe cyclique infini du groupe multiplicatif des nombres complexes. Ses éléments sont

$$\ldots, 2^{-5}, 2^{-4}, \tfrac{1}{8}, \tfrac{1}{4}, \tfrac{1}{2}, 1, 2, 4, 8, 2^4, 2^5, \ldots$$

(ii) le noyau n'est pas trivial. Soit d le plus petit entier positif appartenant au noyau. L'entier d est appelé la *période* (ou l'*ordre*) de a. Si m est un entier tel que $a^m = e, m = ds$ pour un entier s, en vertu de ce qui a été montré au début de ce paragraphe. Remarquons que les éléments

$$e, a, \ldots, a^{d-1}$$

sont distincts. En effet, supposons que $a^r = a^s$, avec $0 \leqq r \leqq d - 1$ et $0 \leqq s \leqq d - 1$, et que, par exemple, $r \leqq s$. Alors $a^{s-r} = e$. Puisque $0 \leqq s - r < d$, on a $s - r = 0$, d'où $r = s$. Nous en déduisons que le groupe cyclique engendré par a est dans ce cas d'ordre d.

Exemple 2. Le groupe multiplicatif $\{1, -1\}$ est cyclique d'ordre 2.

Exemple 3. Les nombres complexes $\{1, i, -1, -i\}$ forment un groupe cyclique d'ordre 4. Le nombre i en est un générateur.

***Théorème* 9.** *Soient G un groupe fini et a un élément de G. L'ordre de a divise l'ordre de G.*

Démonstration. L'ordre du sous-groupe engendré par a est égal à la période de cet élément. On peut donc maintenant appliquer le corollaire du théorème 2, §4.

***Théorème* 10.** *Soit G un groupe cyclique. Tout sous-groupe de G est cyclique.*

Démonstration. Soit a un générateur de G; nous avons donc un homomorphisme surjectif $f: \mathbf{Z} \to G$ tel que $f(n) = a^n$. Soit H un sous-groupe de G. Alors $f^{-1}(H)$ (ensemble des $n \in \mathbf{Z}$ tels que $f(n) \in H$) est un sous-groupe A de $\mathbf{Z}$, et est donc cyclique. Nous savons, en fait, qu'il existe un unique entier positif d tel que $f^{-1}(H)$ est constitué de tous les entiers de la forme md, où $m \in \mathbf{Z}$. Puisque f est surjective, l'application f envoie A sur H tout entier, i.e. tout élément de H est de la forme a^{md}, où m est un entier. Il en résulte que H est cyclique, et qu'en fait a^d en est un générateur.

EXERCICES

1. Montrez qu'un groupe d'ordre 4 est isomorphe à l'un des groupes suivants:

(a) le groupe engendré par deux éléments distincts a et b tels que

$$a^2 = b^2 = e \qquad \text{et} \qquad ab = ba,$$

(b) le groupe G possédant un élément a tel que $G = \{e, a, a^2, a^3\}$ avec $a^4 = e$.

2. Soient m et n deux entiers positifs premiers entre eux. Soient G et G' deux groupes cycliques d'ordre m et n respectivement. Montrez que $G \times G'$ est cyclique et d'ordre mn.

3. Démontrez que tout groupe d'ordre premier est cyclique.

4. Soit G un groupe cyclique d'ordre n; montrez que $a^n = e$, pour tout $a \in G$.

5. Soit G un groupe non commutatif d'ordre 6. Montrez que G est isomorphe à S_3. [*Indication:* montrez que G contient des éléments a et b tels que $a^2 = e$, $b^3 = e$ et $aba = b^2 = b^{-1}$.]

6. Soit $f : G \to G'$ un isomorphisme de groupes. Soit $a \in G$. Montrez que la période de a est la même que la période de $f(a)$.

7. Une *racine de l'unité* dans le corps des nombres complexes est un nombre ω tel que $\omega^n = 1$, pour un entier positif n. Nous disons alors que ω est une racine n-ième de l'unité. Décrivez l'ensemble des racines n-ièmes de l'unité dans $\mathbf{C}$. Montrez que c'est un groupe cyclique d'ordre n.

8. Soient G un groupe cyclique, et $f : G \to G'$ un homomorphisme. Montrez que l'image de G est cyclique.

9. Soit G un groupe cyclique fini d'ordre n. Montrez que pour tout entier positif d divisant n, il existe un sous-groupe d'ordre d.

10. Soit G un groupe cyclique fini d'ordre n. Soit a un générateur de ce groupe. Soit r un entier non nul premier à n. Montrez que a^r est aussi un générateur de G. Montrez que tout générateur de G peut s'écrire sous cette forme.

11. Soient p un nombre premier et G un groupe cyclique d'ordre p. Combien G a-t-il de générateurs?

12. Soient G et X des groupes cycliques d'ordre n. Montrez que $\mathrm{Hom}(G, X)$ est cyclique d'ordre n. [*Indication:* si a est un générateur de G, montrez que, pour tout $x \in X$, il existe un unique homomorphisme $f : G \to X$ tel que $f(a) = x$.]

13. Soit A un groupe abélien, noté additivement, et soit n un entier positif tel que $nx = 0$ pour tout $x \in A$. Supposons que l'on peut écrire $n = rs$, où r et s sont des entiers positifs premiers entre eux. Soit A_r (resp. A_s) l'ensemble de tous les $x \in A$ tels que $rx = 0$ (resp. $sx = 0$). Montrez que tout élément $a \in A$ peut s'écrire d'une manière unique sous la forme $a = b + c$, où $b \in A_r$ et $c \in A_s$.

14. Soient A un groupe abélien additif, et B et C des sous-groupes de A. Soit $B + C$ l'ensemble des sommes $b + c$, où $b \in B$ et $c \in C$. Montrez que $B + C$ est un sous-groupe de A, appelé la somme de B et de C. Définissez la somme d'un nombre fini quelconque de sous-groupes de A de façon analogue.

Nous disons que A est la *somme directe* de B et de C si tout élément x de A s'écrit de manière unique sous la forme $x = b + c$, où $b \in B$ et $c \in C$. On écrit alors $A = B \oplus C$. On procède de la même manière pour un nombre fini quelconque de sous-groupes de A.

15. Montrez qu'un groupe abélien additif A est la somme directe de deux sous-groupes B et C si et seulement si $A = B + C$ et $B \cap C = \{0\}$.

16. Soit A un groupe abélien fini d'ordre n, et soit

$$n = p_1^{r_1} \dots p_s^{r_s}$$

la décomposition de n en facteurs premiers, les p_i étant distincts. Montrez que A est une somme directe $A = A_1 \oplus \cdots \oplus A_s$, où tout élément de A_i a pour ordre un diviseur de $p_i^{r_i}$.

17. Soit G le groupe de l'exercice 8, §1, et soit H le sous-groupe engendré par x. Montrez que G/H est cyclique d'ordre 2.

18. Soit G le groupe de l'exercice 10, §1 et soit H le sous-groupe engendré par x. Montrez que G/H est cyclique d'ordre 2.

CHAPITRE III

Anneaux

Dans ce chapitre, nous axiomatisons les notions d'addition et de multiplication.

§1. *Anneaux*

Un *anneau* A est un ensemble dont les éléments peuvent être additionnés et multipliés (i.e. pour lequel sont données des applications $(x, y) \mapsto x + y$ et $(x, y) \mapsto xy$ de l'ensemble des couples d'éléments de A, dans A), satisfaisant les conditions suivantes:

A 1. *A est un groupe abélien pour l'addition.*

A 2. *Pour tous x, y et z de A, on a*

$$x(y + z) = xy + xz \quad \text{et} \quad (y + z)x = yx + zx.$$

A 3. *Pour tous $x, y, z \in A$, on a $(xy)z = x(yz)$.*

A 4. *Il existe un élément e de A tel que $ex = xe = x$, pour tout $x \in A$.*

Exemple 1. Soit A l'ensemble $\mathbf{Z}$ des entiers; A est un anneau.

Exemple 2. Les nombres rationnels, les nombres réels et les nombres complexes forment tous des anneaux.

Exemple 3. Soit A l'ensemble des fonctions continues réelles définies sur l'intervalle $[0, 1]$. On définit de la manière habituelle la somme et le produit de deux fonctions f et g, par $(f + g)(t) = f(t) + g(t)$ et $(fg)(t) = f(t)g(t)$. Cet ensemble A est un anneau.

Plus généralement, soit E un ensemble non vide, et soit A un anneau. Soit $F(E, A)$ l'ensemble des applications de E dans A. Alors, si l'on définit l'addition et le produit des deux applications f et g par

$$(f + g)(x) = f(x) + g(x) \quad \text{et} \quad (fg)(x) = f(x)g(x),$$

$F(E, A)$ est un anneau. Nous en laissons la vérification, à titre d'exercice simple, au lecteur.

Exemple 4. *L'anneau des endomorphismes.* Soit G un groupe abélien, et soit $\mathrm{End}(G)$ l'ensemble des homomorphismes de G dans lui-même. On dit que $\mathrm{End}(G)$ est l'ensemble des *endomorphismes* de G; ainsi, suivant les notations du chapitre II, §3, $\mathrm{End}(G) = \mathrm{Hom}(G, G)$. Nous savons que $\mathrm{End}(G)$ est un groupe additif. *Si nous prenons comme loi de composition multiplicative sur* $\mathrm{End}(G)$ *la composition ordinaire*

des applications, End(G) *est alors un anneau.* Nous allons le montrer en détail. On sait déjà que A 1 est vérifié. Pour vérifier A 2, considérons des endomorphismes f, g et h de G. On a, pour tout élément x de G,

$$\begin{aligned}(f \circ (g + h))(x) &= f((g + h)(x)) \\ &= f(g(x) + h(x)) = f(g(x)) + f(h(x)). \\ &= f \circ g(x) + f \circ h(x).\end{aligned}$$

Ainsi, $f \circ (g + h) = f \circ g + f \circ h$. La deuxième égalité de A 2 se vérifie de la même façon. On remarque que A 3 n'est pas autre chose que l'associativité de la composition des applications que nous connaissons déjà. L'élément unité de A 4 est l'application identique I. Ainsi nous voyons que End(G) est un anneau.

On dit qu'un anneau A est *commutatif* si $xy = yx$, pour tout $x, y \in A$. Les anneaux des exemples 1, 2 et 3 sont commutatifs. En général, l'anneau de l'exemple 4 ne l'est pas.

Comme dans le cas des groupes, l'élément e d'un anneau A satisfaisant à A 4 est unique et est appelé l'*élément unité* de l'anneau. On le note souvent 1. Remarquons que si $1 = 0$ dans l'anneau A, A est réduit à 0 et est, dans ce cas, appelé *l'anneau nul.*

On peut, dans un anneau A, déduire des axiomes un grand nombre de règles de l'arithmétique ordinaire. Nous allons en donner la liste. On a $0x = 0$ pour tout $x \in A$. *Démonstration:*

$$0x + x = 0x + ex = (0 + e)x = ex = x.$$

Par conséquent, $0x = x$.

On a $(-e)x = -x$, pour tout $x \in A$. *Démonstration:*

$$(-e)x + x = (-e)x + ex = (-e + e)x = 0x = 0.$$

On a $(-e)(-e) = e$. *Démonstration:* en multipliant l'égalité

$$e + (-e) = 0$$

par $-e$, on trouve

$$-e + (-e)(-e) = 0.$$

En ajoutant e aux deux termes de cette égalité, il vient $(-e)(-e) = e$, ce qu'il fallait démontrer.

Nous vous laissons démontrer à titre d'exercice que

$$(-x)y = -xy \quad \text{et que} \quad (-x)(-y) = xy,$$

pour tous $x, y \in A$.

Nous pouvons généraliser A 2, dite *condition de distributivité*, à plusieurs éléments à savoir: si $x, y_1, \dots, y_n$ sont éléments de l'anneau A alors

$$x(y_1 + \cdots + y_n) = xy_1 + \cdots + xy_n.$$

De même, si $x_1, \dots, x_m$ sont des éléments de A,

$$(x_1 + \cdots + x_m)(y_1 + \cdots + y_n) = x_1y_1 + \cdots + x_my_n$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_iy_j.$$

La somme de droite doit être prise pour tous les indices i et j indiqués. Cette formule plus générale peut se démontrer par récurrence, et nous en omettons la démonstration, qui est fastidieuse.

Soit A un anneau. On entend par *sous-anneau* A' de A un sous-ensemble de A tel que l'élément unité de A soit dans A', et, tel que, si x et y sont dans A', alors, $-x$, $x + y$ et xy sont encore dans A'. Il en résulte évidemment que A' est un anneau, les opérations d'addition et de multiplication de A' étant celles de A.

Exemple 5. Les entiers forment un sous-anneau de l'anneau des nombres rationnels, qui est lui-même un sous-anneau de l'anneau des nombres réels.

Exemple 6. Les fonctions réelles différentiables sur **R** forment un sous-anneau de l'anneau des fonctions continues sur **R**.

Soit A un anneau. Il peut arriver qu'il existe des éléments x et y de A tels que $x \neq 0$ et $y \neq 0$ mais $xy = 0$. De tels éléments sont appelés des *diviseurs de zéro*. Un anneau commutatif sans diviseurs de 0, et tel que $1 \neq 0$, est appelé un *anneau intègre*. Un anneau commutatif dont l'ensemble des éléments non nuls forment un groupe pour la multiplication est appelé un *corps*. Remarquons que, dans un corps, on a nécessairement $1 \neq 0$, et qu'un corps ne possède pas de diviseurs de zéro (démonstration?).

Exemple 7. L'ensemble des entiers **Z** est un anneau intègre. Tout corps est un anneau intégre. Nous verrons plus tard que les polynômes sur un corps forment un anneau intègre.

EXERCICES

+ 1. Soit p un entier premier, et soit A le sous-ensemble de l'ensemble des rationnels m/n tels que $n \neq 0$ et que n ne soit pas divisible par p. Montrez que A est un anneau.

+ 2. Montrez que l'anneau des fonctions réelles définies sur $\lfloor 0, 1 \rfloor$ contient des diviseurs de zéro.

3. Soit A un anneau intègre. Si a, b, c sont des éléments de A, a $\neq 0$, et si $ab = ac$, montrez que $b = c$.

4. Soient A un anneau intègre et $a \in A$, $a \neq 0$. Montrez que l'application $x \mapsto ax$ est une application injective de A dans lui-même.

5. Soit A un anneau intègre fini. Montrez que A est un corps. [*Indication:* utilisez l'exercice précédent.]

6. Soit A un anneau tel que $x^2 = x$, pour tout $x \in A$. Montrez que A est commutatif.

7. Soient A un anneau et $x \in A$, Si n est un entier positif, on pose

$$x^n = x \dots x,$$

(n fois). On a alors, pour des entiers positifs m et n.

$$x^{m+n} = x^m x^n.$$

Si $x, y \in A$ et si $xy = yx$, qu'est ce que $(x + y)^n$? (Cf. exercice 2, chapitre I, §2.)

8. Soient A un anneau commutatif et $x \in A$. On dit que x est nilpotent s'il existe un entier positif n tel que $x^n = 0$. Si x et y sont nilpotents, montrez que $x + y$ est nilpotent.

9. Soient A un anneau et U le sous-ensemble de A constitué de tous les éléments x de A tels qu'il existe $y \in A$ tel que $xy = yx = e$. Montrez que U est un groupe. Les éléments de U sont appelés les *unités* de A.

10. Pouvez-vous décrire les unités de l'anneau de l'exercice 1?

11. Soient A un anneau et Z l'ensemble de tous les éléments a de A tels que $ax = xa$, pour tout élément x de A. Montrez que Z est un sous-anneau de A.

12. Soit A l'ensemble des nombres de la forme $a + b\sqrt{2}$, où a et b sont des nombres rationnels. Montrez que A est un anneau, et qu'en fait, A est même un corps.

13. Soit A l'ensemble des nombres de la forme $a + b\sqrt{2}$, où a et b sont des entiers. Montrez que A est un anneau, mais n'est pas un corps.

14. Soit A l'ensemble des nombres de la forme $a + bi$, où a et b sont des entiers et où $i = \sqrt{-1}$. Montrez que A est un anneau. Donnez-en les unités.

15. Soit A l'ensemble des nombres de la forme $a + bi$, où a et b sont rationnels. Montrez que A est un corps.

16. Soient E un ensemble, A un anneau et $f : E \to A$ une application bijective. Pour tous $x, y \in E$, on pose

$$x + y = f^{-1}(f(x) + f(y)) \quad \text{et} \quad xy = f^{-1}(f(x)f(y)).$$

Montrez que cette somme et ce produit définissent une structure d'anneau sur E.

§2. *Idéaux*

Soit A un anneau. On appelle *idéal à gauche* de A toute partie J de A possédant les propriétés suivantes: si $x, y \in J$, $x + y \in J$; $0 \in J$; si $x \in J$ et $a \in A$, $ax \in J$.

En utilisant l'élément négatif $-e$, on voit que si J est un idéal à gauche et que si $x \in J$, alors $-x \in J$ aussi, puisque $-x = (-e)x$. Les éléments d'un idéal à gauche constituent donc un sous-groupe additif de A, et l'on peut tout aussi bien dire qu'un idéal à gauche est un sous-groupe additif J de A tel que, si $x \in J$ et $a \in A$, $ax \in J$.

Remarquons que A est un idéal à gauche, appelé *idéal unité*, et qu'il en est de même du sous-ensemble de A réduit à 0. De la même manière on peut définir les *idéaux à droite* et les *idéaux bilatères*. Un idéal bilatère J est donc par définition un sous-groupe additif de A tel que, si $x \in J$ et $a \in A$, ax et xa appartiennent à J.

Exemple 1. Soit A l'anneau des fonctions réelles continues sur l'intervalle $[0, 1]$. Soit J l'ensemble des fonctions f telles que $f\left(\frac{1}{2}\right) = 0$. Alors J est un idéal (bilatère puisque A est commutatif).

Exemple 2. Si A est l'anneau $\mathbf{Z}$ des entiers, les entiers pairs, i.e. les entiers de la forme $2n$ où $n \in \mathbf{Z}$, forment un idéal. Les entiers impairs en forment-ils un?

Exemple 3. Soient A un anneau et $a \in A$. L'ensemble des éléments xa où $x \in A$, est un idéal à gauche, appelé l'*idéal à gauche principal* engendré par a (vérifiez en détail que c'est un idéal à gauche). On le note (a). Plus généralement, soient $a_1, \ldots, a_n$ des éléments de A. L'ensemble des éléments de la forme

$$x_1 a_1 + \cdots + x_n a_n$$

où $x_i \in A$, est un idéal à gauche noté $(a_1, \ldots, a_n)$. On appelle *générateurs* de cet idéal les éléments $a_1, \ldots, a_n$.

Nous allons donner la démonstration complète de cette assertion pour montrer combien elle est facile, et nous laissons à titre d'exercice les démonstrations d'autres assertions dans les exemples suivants. Si $y_1, \ldots, y_n, x_1, \ldots, x_n \in A$, alors

$$\begin{aligned}(x_1a_1 + \cdots + x_na_n) &+ (y_1a_1 + \cdots + y_na_n)\\ &= x_1a_1 + y_1a_1 + \cdots + x_na_n + y_na_n\\ &= (x_1 + y_1)a_1 + \cdots + (x_n + y_n)a_n.\end{aligned}$$

Si $z \in A$, alors

$$z(x_1a_1 + \cdots + x_na_n) = zx_1a_1 + \cdots + zx_na_n.$$

Enfin,

$$0 = 0a_1 + \cdots + 0a_n.$$

Cela démontre que l'ensemble des éléments $x_1a_1 + \cdots + x_na_n$ où $x_i \in A$ est un idéal à gauche.

Exemple 4. Soit A un anneau. Soient I et J deux idéaux à gauche de A. On désigne par IJ l'ensemble de tous les éléments $x_1y_1 + \cdots + x_ny_n$ où $x_i \in I$ et $y_j \in J$. Il sera facile au lecteur de vérifier que IJ est encore un idéal à gauche. On vérifie aussi que si I, J et K sont des idéaux à gauche, $(IJ)K = I(JK)$.

Exemple 5. Soient I et J des idéaux à gauche. Par définition $I + J$ est le sous-ensemble de A, formé de tous les éléments $x + y$ où $x \in I$ et $y \in J$. L'ensemble $I + J$ est alors un idéal à gauche. Vérifiez-le en détail et montrez en outre que, si I, J et K sont des idéaux à gauche, alors

$$I(J + K) = IJ + IK.$$

Formulez et démontrez les assertions analogues de celles des exemples 4 et 5, pour les idéaux à droite et les idéaux bilatères.

Exemple 6. Soit I un idéal à gauche et notons IA l'ensemble des éléments de la forme $x_1y_1 + \cdots + x_ny_n$ où $x_i \in I$ et $y_i \in A$. Alors IA est un idéal bilatère. Nous en laissons encore la démonstration à titre d'exercice.

Exercices

1. Montrez qu'un corps n'a pas d'autres idéaux que 0 et l'idéal unité.
2. Soit A un anneau commutatif. Si I est un idéal, le produit II est noté I^2. Soient I_1 et I_2 deux idéaux tels que $I_1 + I_2 = A$. Montrez que $I_1^2 + I_2^2 = A$.
3. Soit A l'anneau de l'exercice 1 du paragraphe précédent. Montrez que les éléments m/n de A tels que m soit divisible par p forment un idéal.
4. Soient A un anneau, J_1 et J_2 des idéaux à gauche de A. Montrez que $J_1 \cap J_2$ est un idéal à gauche, et qu'il en est de même pour des idéaux à droite ou pour des idéaux bilatères.
5. Soient A un anneau et $a \in A$. Soit J l'ensemble de tous les éléments x de A tels que $xa = 0$. Montrez que J est un idéal à gauche.
6. Soit A un anneau et soit I un idéal à gauche. Soit J l'ensemble des éléments x de A tels que $xI = 0$ (i.e. $xy = 0$ pour tout $y \in I$). Montrez que J est un idéal bilatère.

7. L'exemple suivant est d'un grand intérêt en analyse. Soit A l'anneau des fonctions indéfiniment différentiables définies, par exemple, sur l'intervalle $-1 < t < 1$. Soit J_n l'ensemble des fonctions $f \in A$ telles que $D^k f(0) = 0$, pour tout entier k tel que $0 \leqq k \leqq n$. Le symbole D désigne ici l'opération de dérivation, de telle sorte que J_n est l'ensemble des fonctions dont toutes les dérivées, au point 0, s'annulent jusqu'à l'ordre n. Montrez que J_n est un idéal de A.

8. Soit A l'anneau des fonctions réelles définies sur l'intervalle $[0, 1]$. Soit E une partie de cette intervalle. Montrez que l'ensemble des fonctions $f \in A$, telles que $f(x) = 0$ pour tout $x \in E$, est un idéal de A.

§3. *Homomorphismes*

Soient A et A' des anneaux. On entend par *homomorphisme d'anneaux* $f : A \longrightarrow A'$ toute application ayant les propriétés suivantes:

$$f(x + y) = f(x) + f(y), \quad f(xy) = f(x)f(y), \quad f(e) = e'$$

pour tous $x, y \in A$ (e et e' sont les éléments unités de A et A' respectivement).

Le *noyau* d'un homomorphisme d'anneau $f: A \to A'$ est son noyau en tant qu'homomorphisme de groupes additifs, i.e. est l'ensemble des éléments x de A tels que $f(x) = 0$. *Exercice:* montrez que le noyau de f est un idéal bilatère de A.

Exemple 1. Soit A l'anneau des fonctions complexes définies sur l'intervalle $[0, 1]$. L'application qui, à toute fonction $f \in A$, associe $f\left(\frac{1}{2}\right)$ est un homomorphisme d'anneaux de A dans $\mathbf{C}$.

Exemple 2. Soit A l'anneau des fonctions réelles définies sur l'intervalle $[0, 1]$. Soit A' l'anneau des fonctions réelles définies sur l'intervalle $[0, \frac{1}{2}]$. Toute fonction $f \in A$ peut être considérée comme une fonction sur l'intervalle $[0, \frac{1}{2}]$, et la considérant comme telle, nous l'appelons la restriction de f à $[0, \frac{1}{2}]$. Soient, plus généralement, un ensemble E et E' une partie de E. Soit A l'anneau des fonctions réelles définies sur E. Pour toute $f \in A$, on note $f \in E'$ la fonction définie sur E' dont la valeur en $x \in E'$ est $f(x)$. On dit alors que $f \in E'$ est la restriction de f à E'. Soit A' l'anneau des fonctions réelles sur E'. Alors l'application

$$f \mapsto f \,|\, E'$$

est un homomorphisme d'anneaux de A dans A'.

Puisque le noyau d'un homomorphisme d'anneaux n'est défini que relativement aux structures de groupes additifs sous-jacents, nous savons qu'un homomorphisme d'anneaux dont le noyau est trivial est injectif.

Soit $f : A \to A'$ un homomorphisme d'anneaux. S'il existe un homomorphisme d'anneaux $g : A' \to A$ tel que $g \circ f$ et $f \circ g$ soient respectivement les applications identiques de A et de A', nous disons que f est un *isomorphisme d'anneaux*.

Comme c'était le cas pour les groupes, si $f: A \to A'$ est un homomorphisme d'anneaux bijectif, f est un isomorphisme d'anneaux. Démonstration laissée au lecteur.

De plus, si $f: A \to A'$ et $g : A' \to A''$ sont des homomorphismes d'anneaux, le

composé $g \circ f: A \to A''$ est aussi un homomorphisme d'anneaux. Preuve également laissée au lecteur.

Nous allons maintenant définir pour les anneaux une notion analogue à celle de groupe quotient.

Soient A un anneau et I un idéal bilatère de A. *Si x et y sont* des éléments de A, on dit que *x est congru à y* modulo I si $x - y$ appartient à I. On écrit cette relation sous la forme

$$x \equiv y \pmod{I}.$$

Il est alors très simple de démontrer les assertions suivantes.

(a) On a $x \equiv x \pmod{I}$.

(b) Si $x \equiv y$ et $y \equiv z \pmod{I}$, alors $x \equiv z$ (mod I).

(c) Si $x \equiv y$, alors $y \equiv x \pmod{I}$.

(d) Si $x \equiv y \pmod{I}$, et si $z \in A$, alors $xz \equiv yz \pmod{I}$, et aussi $zx \equiv zy \pmod{I}$.

(e) Si $x \equiv y$ et $x' \equiv y' \pmod{I}$, alors $xx' \equiv yy' \pmod{I}$. De plus $x + x' \equiv$ $\equiv y + y' \pmod{I}$.

Les démonstrations de ces assertions sont triviales. A titre d'exemple, nous allons donner celle de (e). Les hypothèses signifient que l'on peut écrire

$$x = y + z \quad \text{et} \quad x' = y' + z'$$

où $z, z' \in I$. Alors

$$xx' = (y + z)(y' + z') = yy' + zy' + yz' + zz'.$$

Puisque I est un idéal bilatère, chacun des éléments zy', yz', zz' appartient à I, et il en est donc de même de leur somme. Par sonséquent, $xx' \equiv yy' \pmod{I}$, comme il fallait le démontrer.

Remarque. Cette notion de congruence généralise celle définie pour les entiers au chapitre 1. En effet, si $A = \mathbf{Z}$, la congruence

$$x \equiv y \pmod{n}$$

du chapitre 1, qui signifie que $x - y$ est divisible par n, équivaut à la propriété: $x - y$ appartient à l'idéal engendré par n.

Soit $x \in A$, notons $\bar{x}$ l'ensemble des éléments de A qui sont congrus à $x \pmod{I}$. Si l'on se souvient de la définition d'un groupe quotient, on voit que $\bar{x}$ n'est pas autre chose que la classe additive $x + I$ de x modulo I. Tout élément de cette classe d'équivalence est appelé *représentant* de cette classe.

Soit $\bar{A}$ l'ensemble des classes d'équivalence de A modulo I. En d'autres termes, $\bar{A} = A/I$ est le groupe additif quotient de A par I. Nous savons donc déjà que $\bar{A}$ est un groupe additif. Nous allons maintenant définir sur $\bar{A}$ une multiplication qui en fera un anneau.

Si $\bar{x}$ et $\bar{y}$ sont des classes additives modulo I, définissons leur produit comme la classe de xy, i.e. comme $\overline{xy} + M$. En utilisant la condition (e) ci-dessus, on voit que cette classe est indépendante du choix des représentants x de $\bar{x}$ et de y de $\bar{y}$. Ainsi notre multiplication est bien définie par

$$(\bar{x} + I)(y + I) = (xy + I).$$

Il est maintenant très simple de vérifier que les axiomes des anneaux sont satisfaits. On le sait déjà pour A 1, puisque A/I a été défini comme groupe quotient. Pour ce qui est de A 2, soit $\bar{x}$, $\bar{y}$ et $\bar{z}$ des classes d'équivalence de représentants x, y et z respectivement dans A. Alors $y + z$ est un représentant de $\bar{y} + \bar{z}$ par définition, et $x(y + z)$ est un représentant de $\bar{x}(\bar{y} + \bar{z})$. Mais $x(y + z) = xy + xz$. De plus, xy est un représentant de $\overline{xy}$, et xz est un représentant de $\overline{xz}$. D'où, par définition,

$$\bar{x}(\bar{y} + \bar{z}) = \overline{xy} + \overline{xz}.$$

On montre A 3 de la même manière. Quant à A 4, si e est l'élément unité de A, $\bar{e}$ est l'élément unité de $\bar{A}$ car $ex = x$ est un représentant de $\bar{e}\bar{x}$. Cela prouve que les axiomes sont tous vérifiés.

On appelle $A = A/I$ l'*anneau quotient* de A par I.

Remarquons que l'application $f : A \to A/I$ définie par $f(x) = \bar{x}$, est un homomorphisme d'anneaux de A sur A/I, dont le noyau est I. La vérification est immédiate, et résulte essentiellement des définitions de l'addition et de la multiplication des classes modulo I.

***Théorème* 1.** *Soit $f: A \to B$ un homomorphisme d'anneaux, et soit I son noyau. Pour toute classe C modulo I, l'image $f(C)$ est réduite à un élément de B, et la correspondance*

$$\bar{f}: C \to f(C)$$

est un isomorphisme d'anneaux de A/I sur l'image de f.

Démonstration. Le fait que l'image de f est un sous-anneau de B est laissée comme exercice (exercice 1). Toute classe C est constituée de tous les éléments $x + z$, pour un x donné et pour tous les $z \in I$. Ainsi

$$f(x + z) = f(x) + f(z) = f(x)$$

implique que $f(C)$ est réduit à un élément. On obtient donc bien une application $\bar{f}: C \to f(C)$. Si x, y représentent des classes modulo I, les relations

$$\begin{aligned} f(xy) &= f(x)f(y), \\ f(x + y) &= f(x) + f(y), \\ f(e_A) &= e_B \end{aligned}$$

montrent que $\bar{f}$ est un homomorphisme de A/I dans B. Si $\bar{x} \in A/I$ est tel que $\bar{f}(\bar{x}) = 0$, cela signifie que, pour tout représentant x de $\bar{x}$, on a $f(x) = 0$; par suite, $x \in I$ et $\bar{x} = 0$ (dans A/I). L'application $\bar{f}$ est donc injective. Tout cela démontre notre propos.

Exemple 3. Supposons que $A = \mathbf{Z}$, et que n soit un entier non nul, alors $A/(n) = \mathbf{Z}/(n)$ est appelé l'*anneau des entiers modulo n*. Remarquons que c'est un anneau fini, ayant exactement n éléments (démonstration?). On écrit aussi $\mathbf{Z}/n\mathbf{Z}$ au lieu de $\mathbf{Z}/(n)$.

Exemple 4. Soit A un anneau quelconque, d'élément unité e. Il est facile de montrer par récurrence que l'application $f: \mathbf{Z} \to A$, définie par $f(n) = ne$, est un homomorphisme d'anneaux. Puisque tout homomorphisme d'anneaux de $\mathbf{Z}$

dans A doit être tel que $f(1) = e$, il résulte par récurrence que la seule valeur possible de $f(n)$ est ne, et qu'il n'existe qu'un tel homomorphisme f d'anneaux de $\mathbf{Z}$ dans A. Supposons $A \neq \{0\}$, si bien que $1 \neq 0$; le noyau de f n'est donc pas $\mathbf{Z}$ tout entier, et c'est donc un idéal $n\mathbf{Z}$, pour un certain entier n positif. Il résulte du théoréme 1 que $\mathbf{Z}/n\mathbf{Z}$ est isomorphe à l'image de f. (Remarquez ici que n peut être nul et qu'alors $\mathbf{Z}/n\mathbf{Z} = \mathbf{Z}$ - N.d.T.)

Exercices

1. Soit $f : A \to A'$ un homomorphisme d'anneaux. Montrez que l'image de f est un sous-anneau de A'.

2. Montrez qu'un homomorphisme d'anneaux d'un corps K dans un autre anneau est soit l'application nulle, soit un isomorphisme de K sur son image.

3. Soient n un entier positif et $\mathbf{Z}_n = \mathbf{Z}/n\mathbf{Z}$ l'anneau quotient de Z par n. Montrez que les unités de $\mathbf{Z}_n$ sont précisément les classes résiduelles $\bar{x}$ ayant un représentant $x \neq 0$, premier avec n. (La définition des unités se trouve dans l'exercice 9, §1.)

4. Soit x un entier premier avec n, et soit φ la fonction d'Euler. Montrez que $x^{\varphi(n)} \equiv 1 \pmod{n}$.

5. (a) Soit p un nombre premier. Montrez que dans l'anneau $\mathbf{Z}/(p)$ tout élément non nul possède un inverse pour la multiplication et que les éléments non nuls forment un groupe multiplicatif.

(b) Soit a un entier, $a \not\equiv 0 \pmod{p}$, montrez que $a^{p-1} \equiv 1 \pmod{p}$.

6. Soit F un corps fini ayant q éléments. Montrez que $x^{q-1} = 1$, pour tout élément non nul x de F. Montrez que $x^q = x$, pour tout x de F.

7. Soient n et n' des entiers positifs premiers entre eux. Soient a et b des entiers. Montrez que les congruences

$$x \equiv a \pmod{n} \quad ; \quad x \equiv b \pmod{n'}$$

peuvent être résolues simultanément en x, dans $\mathbf{Z}$.

8. Si p est un nombre premier, et si r est un entier $\geqq 1$, montrez que

$$\varphi(p^r) = (p-1)p^{r-1}.$$

9. Soient A un anneau, I et I' deux idéaux bilatères de A. Supposons que $I \supset I'$. Si $x \in A$, notons $x(I)$ sa classe modulo I. Montrez qu'il existe un (unique) homomorphisme d'anneaux $A/I' \to A/I$ qui envoie $x(I')$ sur $x(I)$.

10. Si n et m sont des entiers $\neq 0$ tels que n divise m, appliquez l'exercice 9 pour obtenir un homomorphisme d'anneaux de $\mathbf{Z}/(m) \to \mathbf{Z}/(n)$.

11. Soient A et A' deux anneaux. Soit $A \times A'$ l'ensemble des couples (x, x'), où $x \in A$ et $x' \in A'$. Montrez que l'on peut faire de $A \times A'$ un anneau en définissant l'addition et la multiplication composante par composante. Quel est, en particulier, l'élément unité de $A \times A'$?

12. (a) Soient m et n des entiers positifs premiers entre eux. Montrez que l'anneau $\mathbf{Z}/(mn)$ est isomorphe à l'anneau $\mathbf{Z}/(n) \times \mathbf{Z}/(m)$ au moyen de l'application.

$$x \ (\mathrm{mod}\ mn) \mapsto (x \bmod m, x \bmod n).$$

(b) Montrez que, si m et n sont des entiers positifs premiers entre eux, $\varphi(mn) = \varphi(m)\varphi(n)$.

13. Soient P l'ensemble des entiers positifs et A l'ensemble des fonctions définies sur P, à

valeur dans un anneau commutatif C. On définit la somme dans A par l'addition ordinaire des fonctions, et le produit par la formule

$$(f * g) = \sum_{xy = m} f(x)g(y),$$

où la somme est prise pour tous les couples possibles d'entiers positifs (x, y) tels que $xy = m$.

(a) Montrez que A est un anneau commutatif dont l'élément unité est la fonction δ telle que $\delta(1) = 1$ et $\delta(x) = 0$, si $x \neq 1$.

(b) Une fonction f est dite multiplicative si $f(mn) = f(m)f(n)$ chaque fois que m et n sont premiers entre eux. Montrez que si f et g sont multiplicatives, $f * g$ l'est aussi.

(c) Soit μ une fonction telle que $\mu(1) = 1$, $\mu(p_1 \dots p_r) = (-1)^r$, si $p_1 \dots p_r$ sont des nombres premiers distincts *et* $\mu(m) = 0$, si m est divisible par p^2 pour un nombre premier p. Montrez que $\mu * \varphi_1 = \delta$ (où φ_1 est la fonction constante de valeur 1). [*Indication:* montrer d'abord que μ est multiplicative, puis démontrer l'assertion pour les puissances des nombres premiers.] La *formule d'inversion de Möbius* de la théorie élémentaire des nombres n'est rien d'autre que la relation $\mu * \varphi_1 * \mathrm{f} = \mathrm{f}$.

14. Soit $f : A \to A'$ un homomorphisme d'anneaux. Soient J' un idéal bilatère de A', et J l'ensemble des éléments x de A tels que $f(x)$ appartient à J. Montrez que J est un idéal bilatère de A.

15. Soient A un anneau commutatif et N l'ensemble des éléments x de A tels que $x^n = 0$, pour un entier positif n. Montrez que N est un idéal.

16. Montrez que, dans l'exercice 15, si $\bar{x}$ est un élément de A/N et s'il existe un entier $n \geqq 1$ tel que $\bar{x}^n = 0$, alors $\bar{x} = 0$.

17. Soit A un anneau commutatif. Un idéal P de A est dit *premier* si $P \neq A$, et si $a, b \in A$ et $ab \in P$ impliquent que $a \in P$ ou $b \in P$. Montrez qu'un idéal de $\mathbf{Z}$, non réduit à 0, est premier si et seulement s'il est engendré par un nombre premier.

18. Soit A un anneau commutatif. Un idéal M de A est dit *maximal* si $M \neq A$ et s'il n'existe pas d'idéal J tel que $A \supset J \supset M$, où $J \neq A$ et $J \neq M$. Montrez que tout idéal maximal est premier.

19. Soit A un anneau commutatif. (a) Montrez qu'un idéal P est premier si et seulement si A/P est intègre. (b) Montrez qu'un idéal M est maximal si et seulement si A/M est un corps.

20. Soit E un ensemble, et soit X une partie de E telle que ni E, ni X ne soient vides. Soit A un anneau et soit $F(E, A)$ l'anneau des applications de E dans A; soit

$$\rho : F(E, A) \to F(X, A)$$

la restriction des applications, i.e. l'application qui, à tout $f \in F(E, A)$, fait correspondre f considérée comme application de X dans A. Montrez que ρ est surjective et décrivez le noyau de ρ.

21. Soit K un corps et soit E un ensemble. Soit x_0 un élément de E. Soit $F(E, K)$ l'anneau des applications de E dans K, et soit J l'ensemble des applications f de $F(E, K)$ telles que $f(x_0) = 0$. Montrez que J est un idéal maximal de $F(E, K)$ et que $F(E, K)/J$ est isomorphe à K.

§4. *Corps des fractions*

Dans les paragraphes précédents, nous avons supposé que le lecteur connaissait le corps des rationnels, de manière à pouvoir donner des exemples de notions plus abstraites. Nous allons maintenant voir comment on peut définir les nombres rationnels à partir des nombres entiers. On sait aussi former des quotients $f/g (g \neq 0)$

de polynômes, et de tels quotients sont appelés des fractions rationnelles. Notre étude s'applique également à cette situation.

Avant de commencer l'étude théorique de cette question, analysons plus attentivement le cas des nombres rationnels. Ce qui est fait à l'école primaire (ou devrait l'être), c'est de donner des règles pour déterminer les cas où deux quotients de nombres entiers sont égaux. C'est bien nécessaire, puisque, par exemple, $\frac{3}{4} = \frac{6}{8}$. La chose importante reste qu'une fraction est déterminée par un couple de nombres —(3, 4) dans l'exemple précédent— mais aussi par d'autres couples —comme (6, 8)—. Si l'on considère tous les couples conduisant à la même fraction comme équivalents, on est sur la bonne voie pour définir les fractions comme classes d'équivalence de couples. On peut ensuite donner des règles d'addition des fractions, et celles que nous allons donner dans le cas général sont exactement les mêmes que celles qui sont (ou devraient être) données à l'école primaire.

Notre étude s'applique à un anneau intègre A quelconque (rappelons-nous qu'intègre signifie que $1 \neq 0$ et que A est commutatif sans diviseurs de 0).

Soient (a, b) et (c, d) des couples d'éléments de A, avec $b \neq 0$, et $d \neq 0$. Nous disons que ces couples sont équivalents si $ad = bc$. Nous affirmons que nous avons là une relation d'équivalence. Revenant à la définition du chapitre I, §5, nous voyons que RE 1 et RE 3 sont ici des évidences. Quant à RE 2, supposons que (a, b) soit équivalent à (c, d) et que (c, d) soit équivalent à (e, f). Par définition $ad = bc$ et $cf = de$. En multipliant la première égalité par f et la seconde par b, on obtient $adf = bcf$ et $bcf = bde$, d'où $adf = bde$, ou encore $daf - dbe = 0$. Donc $d(af - be) = 0$. Puisque A n'a pas de diviseurs de 0, il en résulte que $af - be = 0$, i.e. que $af = be$. Cela veut dire que (a, b) est équivalent à (e, f), et prouve RE 2.

Désignons par a/b la classe d'équivalence de (a, b). Nous allons maintenant indiquer comment additionner et multiplier de telles classes.

Si a/b et c/d sont de telles classes, leur somme est définie par

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

et leur produit par

$$\frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$

Il faut bien sûr maintenant montrer qu'en définissant la somme et le produit comme ci-dessus, le résultat est indépendant du choix des représentants (a, b) et (c, d) des classes données. Nous allons le faire pour la somme. Supposons que $a/b = a'/b'$ et que $c/d = c'/d'$. Il faut montrer que

$$\frac{ad + bc}{bd} = \frac{a'd' + b'c'}{b'd'}.$$

Cela est vrai si et seulement si

$$b'd'(ad + bc) = bd(a'd' + b'c'),$$

ou, en d'autres termes, si

$$(1) \qquad b'd'ad + b'd'bc = bda'd' + bdb'c'.$$

Mais $ab' = a'b$ et $cd' = c'd$, par hypothèse. Cela montre que (1) est vérifié. Nous laissons la démonstration analogue pour le produit en exercice.

Nous affirmons que l'ensemble des fractions a/b, où $b \neq 0$ est un anneau pour les opérations d'addition et de multiplication que l'on vient de définir. Remarquons d'abord que l'on a un élément unité, à savoir $1/1$, où 1 est l'élément unité de A. Il faut maintenant vérifier tous les autres axiomes d'anneau. Cela est fastidieux, mais évident à toutes les étapes de la démonstration. A titre d'exemple, nous allons montrer l'associativité de l'addition. Pour trois fractions a/b, c/d et e/f, on a

$$\left(\frac{a}{b} + \frac{c}{d}\right) + \frac{e}{f} = \frac{ad + bc}{bd} + \frac{e}{f} = \frac{fad + fbc + bde}{bdf}.$$

D'autre part,

$$\frac{a}{b} + \left(\frac{c}{d} + \frac{e}{f}\right) = \frac{a}{b} + \frac{cf + de}{df} = \frac{adf + bcf + bde}{bdf}.$$

Il est alors clair que les expressions de droite de ces deux égalités sont égales, ce qui prouve l'associativité de l'addition. Les autres axiomes sont tout aussi faciles à démontrer et nous nous dispensons de cette routine fastidieuse. Remarquons que notre anneau de fractions est commutatif.

Désignons par K l'anneau de toutes les fractions a/b. Nous affirmons que K est un corps. Pour le voir, il suffit de montrer que tout élément non nul possède un inverse pour la multiplication. Mais l'élément zéro de K est $0/1$, et si $a/b = 0/1$, alors $a = 0$. Par suite tout élément non nul de K peut s'écrire a/b, avec $b = 0$ et $a \neq 0$. Son inverse est b/a, comme on le voit directement en utilisant la définition de la multiplication des fractions.

Remarquons enfin qu'on a une application naturelle de K dans K, à savoir l'application

$$a \mapsto a/1.$$

C'est encore une vérification de routine de voir que cette application est un homomorphisme d'anneaux injectif. Tout homomorphisme d'anneaux injectif sera appelé un *plongement* et ainsi nous voyons que A est plongé dans K de manière naturelle.

Le corps K sera appelé le *corps des fractions* de A. Si $A = \mathbf{Z}$, K est par définition le corps des *nombres rationnels*. Si A est l'anneau des polynômes défini au chapitre suivant, son corps de fractions est appelé le corps des *fractions rationnelles*.

Supposons que A soit un sous-anneau d'un corps L. L'ensemble de tous les éléments de la forme ab^{-1}, où $a, b \in A$ et $b \neq 0$, est, de manière évidente, un corps, et c'est un sous-corps de L. Nous appelons aussi ce corps *le corps de fractions de A dans L*. Il ne peut y avoir aucune confusion de terminologie puisque le corps des fractions de A, défini plus haut, est isomorphe à ce sous-corps de L, au moyen de l'application

$$a/b \mapsto ab^{-1}$$

La vérification en est triviale, et, en raison de ce résultat, l'élément ab^{-1} de L est aussi noté a/b.

Exemple. Soit K un corps et, comme d'habitude, soit $\mathbf{Q}$ le corps des nombres rationnels. Il n'existe pas nécessairement de plongement de $\mathbf{Q}$ dans K (K peut être, par exemple, fini). Cependant s'il existe un plongement de $\mathbf{Q}$ dans K, il n'en existe qu'un. On peut le voir facilement, car tout homomorphisme $f : \mathbf{Q} \to K$ est tel que $f(1) = e$ (élément unité de K). Ainsi, pour $n > 0$, on voit, par récurrence, que $f(n) = ne$, et, par conséquent, que $f(-n) = -ne$. De plus

$$e = f(1) = f(nn^{-1}) = f(n)f(n^{-1}),$$

si bien que $f(n^{-1}) = f(n)^{-1} = (ne)^{-1}$. Pour toute fraction $m/n = mn^{-1}$, où m et n sont des entiers et $n > 0$, on doit donc avoir

$$f(m/n) = (me)/(ne),$$

ce qui montre que f est déterminée de manière unique. Il est d'usage d'identifier $\mathbf{Q}$ à un sous-corps de K, et de considérer tout nombre rationnel comme un élément de K.

Nous allons, pour finir, faire quelques remarques sur le prolongement d'un anneau dans un corps. *Soient A un anneau intègre, et*

$$f : A \to L$$

un plongement de A dans un corps L. Soit K le corps des fractions de A. Alors f admet un unique prolongement en un plongement de K dans L, i.e. *un plongement $f^* : K \to L$ dont la restriction à K est égale à f.*

Pour voir l'unicité, remarquons que si f^* est un prolongement de f, et si

$$f^* : K \to L$$

est un plongement, alors pour tous $a, b \in A$, on doit avoir

$$f\,(^*(a/b)) = f^*(a)/f^*(b) = f(a)/f(b),$$

de sorte que l'effet de f^* sur K est déterminé par celui de f sur A. Réciproquement, on peut définir f^* par la formule

$$f^*(a/b) = f(a)/f(b),$$

et l'on voit immédiatement que la valeur de f^* ne dépend pas du choix du représentant de la fraction a/b, parce que, si $a/b = c/d$, où $a, b, c, d \in A$ et où $bd \neq 0$, alors $f(a)/f(b) = f(c)/f(d)$. On vérifie comme d'habitude que f^* ainsi définie est un homomorphisme, ce qui prouve l'existence du prolongement.

EXERCICES

1. Développez en détail la démonstration de l'existence du prolongement f^* de la fin du paragraphe précédent.

2. Un isomorphisme (d'anneaux) d'un anneau sur lui-même est encore appelé un *automorphisme* de cet anneau. Soient A un anneau intègre et $\sigma : A \to A$ un automorphisme de A. Montrez que σ admet un unique prolongement en un automorphisme du corps des fractions de A.

3. Soient K un corps et $f : Z \to K$ un homomorphisme de l'anneau des entiers dans K. (a) Montrez que le noyau de f est un idéal premier de $\mathbf{Z}$. Si f est un plongement, on dit que K est de *caractéristique* 0. (b) Si $\operatorname{Ker} f \neq \{0\}$, montrez que $\operatorname{Ker} f$ est engendré par un nombre premier p. On dit, dans ce cas, que K est de *caractéristique* p.

4. Soit K un corps de caractéristique p. Montrez que $(x + y)^p = x^p + y^p$, pour tous $x, y \in K$.

5. Soit K un corps fini de caractéristique p. Montrez que l'application $x \mapsto x^p$ est un automorphisme de K.

CHAPITRE IV

Polynômes

Tout au long de ce chapitre, sauf mention expresse du contraire, les corps considérés contiennent un nombre infini d'éléments.

§1. *Algorithme d'Euclide*

Si K est un corps, on appelle *fonction* toute application d'un ensemble dans K.

Soit K un corps. Un *polynôme* sur K, ou à coefficients dans K, est une fonction f de K dans lui-même telle qu'il existe des éléments $a_0, \ldots, a_n$ de K tels que

$$f(t) = a_n t^n + \cdots + a_0$$

pour tout $t \in K$. Soit

$$g(t) = b_m t^m + \cdots + b_0$$

un autre polynôme sur K; on peut alors former la somme $f + g$. Si, par exemple, $n \geqq m$, on peut poser $b_j = 0$, pour $j > m$,

$$g(t) = 0t^n + \cdots + b_m t^m + \cdots + b_0,$$

et l'on peut alors écrire la valeur de la somme $f + g$ sous la forme

$$(f + g)(t) = (a_n + b_n)t^n + \cdots + (a_0 + b_0).$$

La fonction $f + g$ est donc encore un polynôme. Si $c \in K$, alors

$$(cf)(t) = ca_n t^n + \cdots + ca_0.$$

donc cf est un polynôme.

On peut aussi effectuer le produit de deux polynômes, fg, et

$$(fg)(t) = (a_n b_m)t^{n+m} + \cdots + a_0 b_0$$

si bien que fg est encore un polynôme. On écrit, en fait,

$$(fg)(t) = c_{n+m} t^{n+m} + \cdots + c_0,$$

où

$$c_k = \sum_{i=0}^{k} a_i b_{k-i} = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0.$$

Cette expression des c_k provient simplement du regroupement de tous les termes

$$a_i t^i b_{k-i} t^{k-i} = a_i b_{k-i} t^k$$

dans le produit, regroupement qui fournit le terme en t^k. Puisque la somme et le produit de deux polynômes sont encore des polynômes, et puisque la fonction constante 1 est un polynôme, on voit que l'ensemble des polynômes est un sous-anneau de l'anneau des fonctions de K dans lui-même. Cet anneau des polynômes sera noté $K[t]$.

Soit α un élément de K. Nous disons que α est une *racine* de f si $f(\alpha) = 0$.

Théorème 1. *Soit f un polynôme sur K, de la forme*

$$f(t) = a_n t^n + \cdots + a_0.$$

Si $a_0, \ldots, a_n$ ne sont pas tous nuls, f a au plus n racines dans K. Si

$$g(t) = b_n t^n + \cdots + b_0$$

est un autre polynôme tel que $f(t) = g(t)$ pour tous $t \in K$, alors $a_i = b_i$, pour tout i.

Démonstration. Nous avons besoin d'un lemme.

Lemme *Soit f un polynôme sur K, et soit $\alpha \in K$. Il existe alors des éléments $c_0, \ldots, c_n \in K$ tels que*

$$f(t) = c_0 + c_1(t - \alpha) + \cdots + c_n(t - \alpha)^n.$$

Démonstration. Écrivons $t = \alpha + (t - \alpha)$, et remplaçons t par cette valeur dans l'expression de f. Pour tout entier k tel que $1 \leqq k \leqq n$, on a

$$t^k = (\alpha + (t - \alpha))^k = \alpha^k + \cdots + (t - \alpha)^k$$

(ce développement étant celui qui est obtenu par la formule du binôme), et, par conséquent,

$$a_k t^k = a_k \alpha^k + \cdots + a_k (t - \alpha)^k$$

peut s'écrire comme somme de puissances de $(t - \alpha)$ multipliées par des éléments de K. En prenant la somme des $a_k t^k$ pour $k = 0, \ldots, n$, on trouve l'expression de f souhaitée, ce qui démontre le lemme.

Remarquons que, dans le lemme, on a $f(\alpha) = c_0$. Par suite, si $f(\alpha) = 0$, $c_0 = 0$, et l'on peut écrire

$$f(t) = (t - \alpha)h(t),$$

expression dans laquelle on peut encore écrire

$$h(t) = d_1 + d_2(t - \alpha) + \cdots + d_n(t - \alpha)^{n-1}$$

où $d_1, d_2, \ldots, d_n$ sont des éléments de K. Supposons que f a plus de n racines dans K, et supposons, par exemple, que $\alpha_1, \ldots, \alpha_{n+1}$ soient $n + 1$ racines distinctes de f dans K. Soit $\alpha = \alpha_1$, alors $\alpha_i - \alpha_1 \neq 0$ pour $i = 2, \ldots, n + 1$. De

$$0 = f(\alpha_i) = (\alpha_i - \alpha_1)h(\alpha_i),$$

on déduit que $h(\alpha_i) = 0$ pour $i = 2, \ldots, n + 1$. Par récurrence sur n, on voit maintenant que cela est impossible, ce qui prouve que f a au plus n racines dans K.

Supposons enfin que $f(t) = g(t)$ pour tout $t \in K$. Considérons le polynôme

$$f(t) - g(t) = (a_n - b_n)t^n + \cdots + (a_0 - b_0).$$

Tout élément de K est racine de ce polynôme. Par suite et à cause de ce que nous venons de démontrer, nous devons avoir $a_i - b_i = 0$ pour $i = 0, \ldots, n$, soit $a_i = b_i$, ce qui démontre le théorème.

Le théorème 1 montre que si l'on écrit le polynôme f sous la forme

$$f(t) = a_n t^n + \cdots + a_0,$$

où $a_i \in K$, alors les nombres $a_0, \ldots, a_n$ sont déterminés de manière unique. Ces nombres sont appelés les *coefficients* de ce polynôme. Si n est le plus grand entier tel que $a_n \neq 0$, nous dirons que n est le *degré* de f et nous écrirons $n = \deg f$. Nous disons aussi que a_n est le *coefficient dominant* de f. Nous disons que a_0 est le *terme constant* de f. Si f est le polynôme nul, nous convenons que $\deg f = -\infty$. Nous faisons également les conventions suivantes

$$-\infty + (-\infty) = -\infty$$
$$-\infty + a = -\infty, \quad -\infty < a,$$

pour tout entier a, et *aucune autre opération comportant* $-\infty$ *n'est définie.*

C'est pour que le théorème suivant soit vrai sans exception que nous faisons de telles conventions.

Théorème 2. *Soient f et g des polynômes à coefficients dans K. Alors*

$$\deg(fg) = \deg f + \deg g$$

Démonstration. Soient

$$f(t) = a_n t^n + \cdots + a_0 \quad \text{et} \quad g(t) = b_m t^m + \cdots + b_0,$$

où $a_n \neq 0$ et $b_m \neq 0$. D'après la règle de multiplication des polynômes, on voit que

$$f(t)g(t) = a_n b_m t^{n+m} + \text{des termes de plus bas degré,}$$

et $a_n b_m \neq 0$. Par conséquent, $\deg fg = n + m = \deg f + \deg g$. Si f ou g est nul, notre convention pour $-\infty$ fait que notre assertion reste vraie.

Un polynôme de degré 1 est aussi dit *linéaire.*

Corollaire. *L'anneau $K[t]$ n'a pas de diviseurs de zéro, et est par conséquent intègre.*

Démonstration. Si f et g sont des polynômes non nuls, alors $\deg f$ et $\deg g$ sont $\geqq 0$, donc $\deg(fg) \geqq 0$, et $fg \neq 0$, comme on voulait le montrer.

On trouve, dans le théorème suivant, l'algorithme d'Euclide, ou division telle qu'on l'enseigne à l'école primaire. Cet algorithme d'Euclide est l'analogue de celui donné pour les entiers.

Théorème 3. *Soient f et g des polynômes sur le corps K,* i.e. *des polynômes*

de $K[t]$, *et supposons que* $\deg g \geqq 0$. *Il existe alors des polynômes* q *et* r *de* $K[t]$ *tels que*

$$f(t) = q(t)g(t) + r(t),$$

avec $\deg r < \deg g$. *Les polynômes* q *et* r *sont déterminés de manière unique par ces conditions.*

Démonstration. Soit $m = \deg g \geqq 0$. Posons

$$f(t) = a_n t^n + \cdots + a_0,$$
$$g(t) = b_m t^m + \cdots + b_0,$$

où $b_m \neq 0$. Si $n < m$, on prend $q = 0$ et $r = f$. Si $n \geqq m$, posons

$$f_1(t) = f(t) - a_n b_m^{-1} t^{n-m} g(t)$$

(C'est là la première étape du processus de division.) On a alors $\deg f_1 < \deg f$. On continue de la même manière, ou, plus formellement, par récurrence sur n, et l'on peut trouver des polynômes q_1 et r tels que

$$f_1 = q_1 g + r,$$

où $\deg r < \deg g$. Alors

$$\begin{aligned} f(t) &= a_n b_m^{-1} t^{n-m} g(t) + f_1(t) \\ &= a_n b_m^{-1} t^{n-m} g(t) + q_1(t)g(t) + r(t) \\ &= (a_n b_m^{-1} t^{n-m} + q_1)g(t) + r(t), \end{aligned}$$

et l'on a, par conséquent, exprimé notre polynôme sous la forme voulue. Pour démontrer l'unicité, supposons que

$$f = q_1 g + r_1 = q_2 g + r_2,$$

où $\deg r_1 < \deg g$ et $\deg r_2 < \deg g$. Alors

$$(q_1 - q_2)g = r_2 - r_1,$$

Le degré du terme de gauche est $\geqq \deg g$, sinon ce terme est nul. Le degré du terme de droite est $< \deg g$, sinon ce terme est nul. Par conséquent la seule possibilité est que ces deux termes soient nuls, c'est-à-dire que

$$q_1 = q_2 \quad \text{et} \quad r_1 = r_2,$$

comme il fallait le démontrer.

A l'aide de l'algorithme d'Euclide, on peut redémontrer une propriété déjà établie, par d'autres arguments.

***Corollaire* 1.** *Soit* f *un polynôme non nul de* $K[t]$. *Soit* $\alpha \in K$ *tel que* $f(\alpha) = 0$. *Il existe alors un polynôme* $q(t)$ *de* $K[t]$ *tel que*

$$f(t) = (t - \alpha)q(t).$$

Démonstration. On peut écrire

$$f(t) = q(t)(t - \alpha) + r(t),$$

où $\deg r < \deg(t - \alpha)$. Mais $\deg(t - \alpha) = 1$, donc r est constant. Puisque

$$0 = f(\alpha) = q(\alpha)(\alpha - \alpha) + r(\alpha) = r(\alpha),$$

$r = 0$, comme souhaité.

Corollaire 2. *Soit K un corps tel que tout polynôme, non constant, de $K[t]$ possède une racine dans K. Soit f un tel polynôme. Il existe alors des éléments $\alpha_1, \ldots, \alpha_n$ et c de K tels que*

$$f(t) = c(t - \alpha_1)(t - \alpha_2) \cdots (t - \alpha_n).$$

Démonstration. Remarquons que, dans le corollaire 1, $\deg q = \deg f - 1$. Faisons $\alpha = \alpha_1$ dans le corollaire 1. Par hypothèse, si q n'est pas constant, on peut trouver une racine α_2 de q; on peut donc écrire

$$f(t) = q_2(t)(t - \alpha_1)(t - \alpha_2).$$

En procédant par récurrence, on continue de la sorte jusqu'à ce que q_{n+1} soit constant.

Un corps K possédant la propriété du corollaire 2, à savoir que tout polynôme non constant sur K y possède une racine, est dit *algébriquement clos.* Nous démontrons plus loin dans ce livre que le corps des nombres complexes est algébriquement clos.

Exercices

1. Mettez, dans chacun des cas suivants, f sous la forme $f = qg + r$, avec $\deg r < \deg g$.

(a) $f(t) = t^2 - 2t + 1$, $g(t) = t - 1$
(b) $f(t) = t^3 + t - 1$, $g(t) = t^2 + 1$
(c) $f(t) = t^3 + t$, $g(t) = t$
(d) $f(t) = t^3 - 1$, $g(t) = t - 1$.

2. Montrez que si $f(t)$ est un polynôme à coefficients entiers et $g(t)$ un polynôme à coefficients entiers de coefficient dominant égal à 1, on peut exprimer f sous la forme $qg + r$, où $\deg r <$ $< \deg g$. Les polynômes q et r étant à coefficients entiers.

3. Montrez, en utilisant le théorème des valeurs intermédiaires, que tout polynôme à coefficients réels de degré impair possède une racine réelle.

4. Soient $f(t) = t^n + \cdots + a_0$ un polynôme à coefficients complexes, de degré n, et α une racine de f. Montrez que $|\alpha| \leqq n \cdot \max_i |a_i|$. [*Indication:* écrire

$$-\alpha^n = a_{n-1}\alpha^{n-1} + \cdots + a_0.$$

Si $|\alpha| > n \cdot \max_i |a_i|$, diviser par α^n, prendre la valeur absolue et utiliser une approximation simple pour arriver à une contradiction.]

§2. *Plus grand commun diviseur*

Ayant l'algorithme d'Euclide à notre disposition, nous pouvons maintenant développer la théorie de la divisibilité, exactement comme nous l'avons fait, au chapitre 1, pour les entiers.

Théorème 4. *Soit J un idéal de $K[t]$. Il existe un polynôme g qui est un générateur de J.*

Démonstration. Supposons que J soit un idéal non nul, et soit g un polynôme non nul de J, de degré minimum. Nous affirmons que g est un générateur de J. Soit en effet f un élément quelconque de J; on peut trouver, au moyen de l'algorithme d'Euclide, des polynômes q et r tels que $f = qg + r$, avec $\deg r < \deg g$. Le polynôme $r = f - qg$ est aussi dans J, par définition d'un idéal. Puisque $\deg r < \deg g$, on doit avoir $r = 0$. Par suite, $f = qg$ et g est un générateur de J, comme souhaité.

Remarque. Soit g_1 un générateur non nul d'un idéal J, et soit g_2 un autre générateur. Il existe alors un polynôme q, tel que $g_1 = qg_2$. Puisque

$$\deg g_1 = \deg q + \deg g_2,$$

il s'ensuit que $\deg g_2 \leqq \deg g_1$. Par raison de symétrie, *on doit avoir*

$$\deg g_1 = \deg g_2$$

Donc q est constant. On peut écrire

$$g_1 = cg_2$$

où c est un nombre. Posons

$$g_2(t) = a_n t^n + \cdots + a_0$$

où $a_n \neq 0$. Prenons $b = a_n^{-1}$: bg_2 est alors aussi un générateur de J, et son coefficient dominant est 1. On peut donc toujours trouver un générateur d'un idéal (non nul) de coefficient dominant 1. Il est, de plus, clair que ce générateur est déterminé de manière unique.

Soient f et g des polynômes non nuls. Nous disons que g *divise* f et nous écrivons $g \mid f$, s'il existe un polynôme q tel que $f = gq$. Soient f_1 et f_2 des polynômes non nuls. Nous disons qu'un polynôme g est un *plus grand commun diviseur* de f_1 et de f_2, si g divise f_1 et f_2 et si, de plus, tout polynôme h divisant f_1 et f_2 divise g.

Théorème 5. *Soient f_1 et f_2 des polynômes non-nuls de $K[t]$. Soit g un générateur de l'idéal engendré par f_1 et f_2. Alors g est un plus grand diviseur de f_1 et de f_2.*

Démonstration. Puisque f_1 appartient à l'idéal engendré par f_1 et f_2, il existe un polynôme q_1 tel que $f_1 = q_1 g$, donc g divise f_1. De la même manière g divise f_2. Soit h un polynôme divisant à la fois f_1 et f_2. Écrivons

$$f_1 = h_1 h \quad \text{et} \quad h_2 h,$$

où h_1 et h_2 sont des polynômes. Puisque g appartient à l'idéal engendré par f_1 et f_2, il existe des polynômes g_1 et g_2 tels que $g = g_1f_1 + g_2f_2$, par suite

$$g = g_1h_1h + g_2h_2h = (g_1h_1 + g_2h_2)h.$$

Par suite h divise g, et notre théorème est démontré.

Remarque 1. Le plus grand commun diviseur est déterminé à une constante multiplicative non nulle près. Si l'on choisit le plus grand commun diviseur à coefficient dominant 1, il est déterminé de manière unique.

Remarque 2. On peut donner exactement la même démonstration pour plus de deux polynômes. Par exemple, si $f_1, ..., f_n$ sont des polynômes non nuls, et si g est un générateur de l'idéal engendré par $f_1, ..., f_n$, alors g est un plus grand diviseur commun de $f_1, ..., f_n$.

Des polynômes $f_1, ..., f_n$ dont le plus grand commun diviseur est 1 sont dits *premiers entre eux* (ou *étrangers* - N.d.T.).

EXERCICES

1. Montrez que $t^n - 1$ est divisible par $t - 1$.
2. Montrez que $t^4 + 4$ peut se mettre sous la forme d'un produit de deux polynômes de degré 2, à coefficients entiers.
3. Trouvez le quotient de $t^n + 1$ par $t + 1$, pour n impair.

§3. *Unicité de la décomposition*

Un polynôme p de $K[t]$ est dit *irréductible* (sur K) s'il est de degré $\geqq 1$, et si, étant donnée une décomposition de p sous la forme $p = fg$, où $f, g \in K[t]$, $\deg f$ ou $\deg g = 0$ (i.e. l'un des polynômes f ou g est une constante). Ainsi, à une constante multiplicative non nulle près, les seuls diviseurs de p sont p lui-même et 1.

Exemple 1. Les seuls polynômes irréductibles sur les corps des nombres complexes sont les polynômes de degré 1, i.e. les produits par une constante non nulle de polynômes du type $t - \alpha$, où $\alpha \in \mathbf{C}$.

Exemple 2. Le polynôme $t^2 + 1$ est irréductible sur $\mathbf{R}$.

Théorème 6. *Tout polynôme de $K[t]$ de degré $\geqq 1$ peut s'exprimer comme produit $p_1 \cdots p_m$ de polynômes irréductibles. Dans un tel produit, les polynômes $p_1, ..., p_m$ sont déterminés de manière unique, à l'ordre près, et à une constante multiplicative non nulle près.*

Démonstration. Montrons d'abord l'existence d'une telle décomposition en un produit de polynômes irréductibles. Soit $f \in K[t]$ de degré $\geqq 1$. Si f est irréductible, il n'y a rien à démontrer; sinon, on peut écrire

$$f = gh,$$

où $\deg g < \deg f$ et $\deg h < \deg. f$. Par hypothèse de récurrence, on peut exprimer g et h comme produits de polynômes irréductibles, et, par conséquent, $f = gh$ peut aussi s'exprimer comme un tel produit.

Il nous faut maintenant démontrer l'unicité. Nous avons besoin du lemme suivant.

Lemme. *Soit p un polynôme irréductible de $K[t]$. Soient f et g des polynômes non nuls de $K[t]$, et supposons que p divise fg. Alors p divise f ou p divise g.*

Démonstration. Supposons que p ne divise pas f. Le plus grand commun diviseur de p et de f est alors 1, et il existe des polynômes h_1 et h_2 de $K[t]$ tels que

$$1 = h_1 p + h_2 f$$

(On utilise le théorème 5.) Multiplions cette égalité par g, on obtient

$$g = gh_1 p + h_2 fg.$$

Mais il existe h_3 tel que $fg = ph_3$, d'où

$$g = (gh_1 + h_2 h_3)p,$$

et p divise g, ce qu'il fallait démontrer.

Nous appliquons ce lemme lorsque p divise un produit de polynômes irréductibles $q_1, \ldots, q_s$. Dans ce cas, p divise q_1 ou p divise $q_2, \ldots, q_s$. Il existe donc une constante c telle que $p = cq_1$, sinon p divise $q_2, \ldots, q_s$. Dans ce dernier cas, on peut procéder par récurrence et en conclure que, dans tous les cas, il existe un i tel que p et q_i ne diffère que d'une constante multiplicative.

Supposons maintenant que nous avons deux produits de polynômes irréductibles

$$p_1 \cdots p_r = q_1 \cdots q_s.$$

Après avoir éventuellement ré-indexer les q_i, on peut supposer que $p_1 = c_1 q_1$, où c_1 est une constante. En simplifiant par q_1, on obtient

$$c_1 p_2 \cdots p_r = q_2 \cdots q_s.$$

En itérant ce procédé, on conclut qu'il existe des constantes c_i telles que $p_i = c_i q_i$, après une permutation éventuelle des q_i. Cela démontre l'unicité.

***Corollaire* 1.** *Soit f un polynôme de $K[t]$ de degré $\geqq 1$. Le polynôme f possède une décomposition $f = cp_1 \ldots p_s$, où $p_1 \ldots p_s$ sont des polynômes irréductibles, de coefficients dominants 1, déterminés de manière unique à une permutation près.*

***Corollaire* 2.** *Soit K un corps algébriquement clos. Soit f un polynôme de $K[t]$ de degré $\geqq 1$. Le polynôme f possède alors une décomposition sous la forme*

$$f(t) = c(t - \alpha_1) \cdots (t - \alpha_n),$$

où $\alpha_i \in K$ et $c \in K$. Les facteurs $t - \alpha_i$ sont déterminés de manière unique à une permutation près.

Nous avons, la plupart du temps, affaire à des polynômes ayant 1 pour coefficient dominant. Soit f un tel polynôme de degré $\geqq 1$. Soient $p_1, \ldots, p_r$ les polynômes irréductibles distincts (de coefficients dominants 1) intervenant dans la décomposition de f. On peut alors exprimer f comme un produit de la forme

$$f = p_1^{i_1} \ldots p_r^{i_r},$$

où $i_1, \ldots, i_r$ sont des entiers positifs, déterminés de manière unique par $p_1, \ldots, p_r$. Nous disons que cette décomposition est la décomposition normalisée de f. En particulier, sur un corps algébriquement clos,

$$f(t) = (t - \alpha_1)^{i_1} \cdots (t - \alpha_r)^{i_r}.$$

Un polynôme ayant 1 pour coefficient dominant est parfois appelé *monique* (ou *unitaire* - N.d.T.).

Si p est irréductible, et si $f = p^m g$, où p ne divise pas g et où m est un entier positif ou nul, on dira que m est la *multiplicité* de p dans f (on pose $p^0 = 1$ par définition). On note cette multiplicité $\operatorname{ord}_p f$, et on l'appelle encore ordre de f en p.

Si α est une racine de f, et si

$$f(t) = (t - \alpha)^m g(t),$$

où $g(\alpha) \neq 0$, alors $t - \alpha$ ne divise pas $g(t)$, et m est la multiplicité de $t - \alpha$ dans f. On dit aussi que m est la *multiplicité* de α dans f (ou l'*ordre* de la racine α dans f - N.d.T.)

Il y a un critére simple, utilisant la dérivée, pour vérifier si $m > 1$.

Soit $f(t) = a_n t^n + \cdots + a_0$ un polynôme. On définit sa *dérivée* (*formelle*) comme suit

$$Df(t) = f'(t) = n a_n t^{n-1} + (n - 1) a_{n-1} t^{n-2} + \cdots + a_1.$$

Nous pouvons alors affirmer ce qui suit, en laissant au lecteur le soin d'en faire la démonstration.

Si f et g sont des polynômes, alors

$$(f + g)' = f' + g'$$

et

$$(fg)' = f'g + fg'.$$

Si c est une constante, $(cf)' = cf'$.

Si $f(t) = h(t)^m$ pour un entier $m \geqq 1$, $f'(t) = m h(t)^{m-1} h'(t)$ (démontrer cette dernière assertion par récurrence).

Théorème 7. *Soit K un corps. Soit f un polynôme sur K de degré* 1, *et soit α une racine de f dans K. La multiplicité de α dans f est alors > 1 si et seulement si $f'(\alpha) = 0$.*

Démonstration. Supposons que $f(t) = (t - \alpha)^m g(t)$, où $m > 1$. En prenant la dérivée, on trouve

$$f'(t) = m(t - \alpha)^{m-1}g(t) + (t - \alpha)^m g'(t).$$

En remplaçant t par α dans cette expression, on voit que $f'(\alpha) = 0$ puisque $m - 1 \geqq 1$. Réciproquement, supposons que

$$f(t) = (t - \alpha)^m g(t),$$

et que $g(\alpha) \neq 0$, si bien que m est l'ordre de α. Si $m = 1$.

$$f'(t) = g(t) + (t - \alpha)g'(t),$$

de telle sorte que $f'(\alpha) = g(\alpha) \neq 0$. Cela démontre notre théorème.

Exercices

1. Soit f un polynôme de degré 2 sur un corps K. Montrez que, soit f est irréductible sur K, soit f se décompose en deux facteurs du premier degré sur K.

2. Soit f un polynôme de degré 3 sur K. Montrez que si f n'est pas irréductible sur K, f possède une racine dans K.

3. Soit $f(t)$ un polynôme irréductible à coefficients réels ayant 1 pour coefficient dominant. Supposons que $\deg f = 2$. Montrez que $f(t)$ peut s'écrire sous la forme

$$f(t) = (t - a)^2 + b^2,$$

pour $a, b \in \mathbf{R}$, $b \neq 0$. Montrez réciproquement qu'un tel polynôme est irréductible.

4. Soit f un polynôme à coefficients complexes par exemple

$$f(t) = \alpha_n t^n + \cdots + \alpha_0.$$

Le conjugué de f sera par définition

$$f(t) = \bar{\alpha}_n t^n + \cdots + \bar{\alpha}_0,$$

où on a pris le conjugué de chaque coefficient de f. Montrez que *si* f et g sont dans $\mathbf{C}[t]$,

$$\overline{(f + g)} = \bar{f} + \bar{g}, \quad \overline{(fg)} = \bar{f}\bar{g},$$

et que si $\beta \in \mathbf{C}$, alors $\overline{(\beta f)} = \bar{\beta}\bar{f}$.

5. Soit $f(t)$ un polynôme à coefficients réels. Soit α une racine complexe non réelle de f. Montrez que $\bar{\alpha}$ est aussi racine de f.

6. Les données étant celles de l'exercice 5, montrez que l'ordre de α dans f est le même que l'ordre de $\bar{\alpha}$.

7. Montrez que les polynômes suivants n'ont pas de racines multiples dans $\mathbf{C}$.

(a) $t^4 + t$; (*b*) $t^5 - 5t + 1$

(c) Tout polynôme de la forme $t^2 + bt + c$, où b et c sont tels que $b^2 - 4c \neq 0$.

8. Montrez que le polynôme $t^n - 1$ n'a pas de racines multiples dans $\mathbf{C}$. Pouvez-vous en déterminer toutes les racines et donner sa décomposition en facteurs de degré 1?

9. Soit K un sous-corps de $\mathbf{C}$, et soit $\alpha \in \mathbf{C}$. Soit J l'ensemble de tous les polynômes $f(t)$ de $K[t]$ tels que $f(\alpha) = 0$. Montrez que J est un idéal. Montrez que si J est un idéal non nul, le générateur monique de J est irréductible.

10. Soient f et g des polynômes écrits sous la forme

$$f = p_1^{i_1} \ldots p_r^{i_r} \quad \text{et} \quad g = p_1^{j_1} \ldots p_r^{j_r},$$

où i_ν et j_ν sont des entiers positifs ou nuls, et où $p_1, \ldots, p_r$ sont des polynômes irréductibles distincts.

(a) Montrez que le plus grand commun diviseur de f et de g peut s'exprimer comme un produit $p_1^{k_1}, \ldots, p_r^{k_r}$, où $k_1, \ldots, k_r$ sont des entiers $\geqq 0$. Exprimez k_ν en fonction de i_ν et de j_ν.

(b) Définissez le plus petit commun multiple de plusieurs polynômes et exprimez le plus petit commun multiple de f et de g comme produit $p_1^{k_1} \ldots p_r^{k_r}$, où les entiers k_ν sont $\geqq 0$. Exprimez k_ν en fonction de i_ν et de j_ν.

11 Trouvez le plus grand commun diviseur et le plus petit commun multiples de chacun des couples suivants:

(a) $(t-2)^3(t-3)^4(t-i)$ et $(t-1)(t-2)(t-3)^3$

(b) $(t^2+1)(t^2-1)$ et $(t+i)^3(t^3-1)$.

12. Soit K un corps, $A = K[t]$ l'anneau des polynômes à coefficients dans K, et F le corps des fractions de A, i.e. le corps des fractions rationnelles. Soit $\alpha \in K$, Soit A_α l'ensemble des fractions rationnelles qui peuvent s'écrire f/g pour des polynômes g et f tels que $g(\alpha) \neq 0$. Montrez que A_α est un anneau. Si φ est une fraction rationnelle, et si $\varphi = f/g$, où $g(\alpha) \neq 0$, on pose $\varphi(\alpha) = f(\alpha)/g(\alpha)$. Montrez que $\varphi(\alpha)$ ne dépend pas du choix de la représentation de φ comme quotient f/g. Montrez que l'application $\varphi \mapsto \varphi(\alpha)$ est un homomorphisme d'anneaux de A_α dans K. Montrez que le noyau de cet homomorphisme d'anneaux est formé de toutes les fractions rationnelles f/g telles que $g(\alpha) \neq 0$ et $f(\alpha) = 0$. Si l'on note M_α ce noyau, montrez que M_α est un idéal premier de A_α.

13. Soit A un anneau commutatif, soit K un sous-corps de A (i.e. un sous-anneau qui est aussi un corps). Soit $k \in K[t]$ un polynôme à coefficients dans K,

$$f(t) = a_n t^n + \cdots + a_0.$$

Soit $b \in A$. Posons

$$f(b) = a_n b^n + \cdots + a_0.$$

Montrez que l'application $f \mapsto f(b)$ est un homomorphisme d'anneaux de $K[t]$ dans A.

14. Soit R une fraction rationnelle à coefficients dans le corps K, exprimée comme quotient de polynômes, $R = g/f$. Définissons la dérivée de R comme étant

$$R' = \frac{fg' - gf'}{f^2},$$

où le signe «prime» désigne la dérivée formelle du polynôme, comme ci-dessus.

(a) Montrez que la dérivée de R ne dépend pas de son expression comme quotient de polynômes, i.e. si $R = g_1/f_1$, alors

$$\frac{fg' - gf'}{f^2} = \frac{f_1 g_1' - g_1 f_1'}{f_1^2}.$$

(b) Montrez que la dérivée des fractions rationnelles satisfait aux propriétés usuelles, i.e. qu'on a, pour des fractions rationnelles R_1 et R_2,

$$(R_1 + R_2)' = R_1' + R_2' \quad \text{et} \quad (R_1 R_2)' = R_1 R_2' + R_1' R_2.$$

(c) Soient $\alpha_1, \ldots, \alpha_n$ et $a_1, \ldots, a_n$ des éléments de K tels que

$$\frac{1}{(t-\alpha_1)\cdots(t-\alpha_n)} = \frac{a_1}{t-\alpha_1} + \cdots + \frac{a_n}{t-\alpha_n}.$$

Soit $f(t) = (t-\alpha_1)\cdots(t-\alpha_n)$; supposons que $\alpha_1, \ldots, \alpha_n$ sont distincts.

$$\text{Montrez que } a_1 = \frac{1}{(\alpha_1-\alpha_2)\cdots(\alpha_1-\alpha_n)} = \frac{1}{f'(\alpha_1)}.$$

§4. *Décomposition en éléments simples*

Nous avons démontré au paragraphe précédent qu'on peut exprimer un polynôme comme produit de puissances de polynômes irréductibles d'une façon unique à une permutation près des facteurs. La même chose reste vraie, si on admet l'utilisation d'exposants négatifs. Soit $R = g/f$ une fraction rationnelle, exprimée comme quotient des polynômes g et f, où $f \neq 0$. Supposons que $R \neq 0$. Si g et f ne sont pas premiers entre eux, on peut simplifier par leur plus grand commun diviseur pour obtenir R comme quotient de deux polynômes premiers entre eux. En mettant en facteurs leurs coefficients dominants, on peut écrire

$$R = c\frac{g_1}{f_1},$$

où f_1 et g_1 ont 1 pour coefficient dominant. Alors f_1, g_1 et c sont déterminés de manière unique; supposons que

$$cg_1/f_1 = c_2g_2/f_2,$$

pour des constantes c et c_2 et pour des couples de polynômes étrangers f_1, g_1 et f_2, g_2 de coefficients dominants égaux à 1. On a alors

$$cg_1f_2 = c_2g_2f_1.$$

D'après l'unicité de la décomposition des polynômes, nous arrivons à la conclusion que $g_1 = g_2$ et $f_1 = f_2$, de sorte que $c = c_2$.

Si nous décomposons maintenant f_1 et f_1 en produits de puissances de polynômes irréductibles, nous obtenons l'unique décomposition possible de R. Tout cela est parfaitement analogue à la décomposition obtenue pour les nombres rationnels au chapitre I, §4.

Nous voulons maintenant décomposer une fraction rationnelle en somme de fractions rationnelles, telles que le dénominateur de chacune d'elles soit une puissance d'un polynôme irréductible. Une telle décomposition est appelée *décomposition en éléments simples*. Commençons par un lemme qui nous permet de raisonner par récurrence.

Lemme. *Soient f_1 et f_2 deux polynômes non nuls et étrangers à coefficients dans un corps K. Il existe alors des polynômes h_1 et h_2, à coefficients dans K, tels que*

$$\frac{1}{f_1 f_2} = \frac{h_1}{f_1} + \frac{h_2}{f_2}.$$

Démonstration. Puisque f_1 et f_2 sont étrangers, il existe des polynômes h_1 et h_2 tels que

$$h_2 f_1 + h_1 f_2 = 1.$$

En divisant les deux côtés de cette égalité par $f_1 f_2$, on obtient le résultat.

Théorème 8. *Toute fraction rationnelle peut s'écrire sous la forme*

$$R = \frac{h_1}{p_1^{i_1}} + \cdots + \frac{h_n}{p_n^{i_n}} + h,$$

où $p_1, \ldots, p_n$ sont des polynômes irréductibles distincts de coefficients dominants égaux à 1; $i_1, \ldots, i_n$ sont des entiers $\geqq 0$; $h_1, \ldots, h_n$, h sont des polynômes satisfaisant à

$$\deg h_\nu < \deg p_\nu^i \quad \text{et} \quad p \nmid h_\nu$$

pour $\nu = 1, \ldots, n$. Dans une telle expression les entiers i_ν et les polynômes h_ν, $h(\nu = 1, \ldots, n)$ sont déterminés de manière unique.

Démonstration. Nous allons d'abord démontrer l'existence de l'expression donnée dans notre théorème. Soit $R = g/f$, où f est un polynôme non nul, et posons

$$f = p_1^{i_1} \cdots p_n^{i_n}$$

où $p_1, \ldots, p_n$ sont des polynômes irréductibles distincts et $i_1, \ldots, i_n$ des entiers $\geqq 0$. D'après le lemme, il existe des polynômes g_1, g_1^* tels que

$$\frac{1}{f} = \frac{g_1}{p_1^{i_1}} + \frac{g_1^*}{p_2^{i_2} \cdots p_n^{i_n}}$$

et par hypothèse de récurrence, il existe des polynômes $g_2, \ldots, g_n$ tels que

$$\frac{g_1^*}{p_2^{i_2} \cdots p_n^{i_n}} = \frac{g_2}{p_2^{i_2}} + \cdots + \frac{g_n}{p_n^{i_n}}$$

En multipliant par g, il vient

$$\frac{g}{f} = \frac{g g_1}{p_1^{i_1}} + \cdots + \frac{g g_n}{p_n^{i_n}}.$$

Nous pouvons diviser $g g_\nu$ par $p_\nu^{i_\nu}$ en utilisant l'algorithme d'Euclide pour $\nu = 1, \ldots, n$ et nous obtenons

$$g g_\nu = q_\nu p_\nu^{i_\nu} + h_\nu, \quad \deg h_\nu < \deg p_\nu^{i_\nu}.$$

On obtient de cette façon l'expression souhaitée de g/f, où $h = q_1 + \cdots + q_n$.

Démontrons maintenant l'unicité de cette décomposition. (On peut supposer que les polynômes irréductibles $p_1, \dots, p_n$ sont les mêmes des deux côtés, en posant $i_\nu = 0$ pour certains i_ν, si nécessaire.) Il existe alors des polynômes φ, ψ tels que $\psi \neq 0$ et $p_1 \nmid \psi$, et pour lesquels nous pouvons écrire

$$\frac{h_1}{p_1^{i_1}} - \frac{\bar{h}_1}{p_1^{j_1}} = \frac{\varphi}{\psi} .$$

Supposons que, par exemple, $i_1 \leqq j_1$. On a alors

$$\frac{h_1 p_1^{j_1 - i_1} - \bar{h}_1}{p_1^{j_1}} = \frac{\varphi}{\psi} .$$

Puisque ψ n'est pas divisible par p_1, il résulte de l'unicité de la décomposition que $p_1^{j_1}$ divise $h_1 p_1^{j_1 - i_1} - \bar{h}_1$. Si $j_1 \neq i_1$, alors $p_1 \mid \bar{h}_1$, contrairement aux hypothèses du théorème. D'où $j_1 = i_1$. Mais, puisque ψ n'est pas divisible par p_1, il s'ensuit maintenant que $p_1^{j_1}$ divise $h_1 - \bar{h}_1$. Or, par hypothèse,

$$\deg (h_1 - \bar{h}_1) < \deg p_1^{j_1}$$

D'où $h_1 - \bar{h}_1 = 0$, et par suite $h_1 = \bar{h}_1$. Nous en concluons donc que

$$\frac{h_2}{p_2^{i_2}} + \cdots + \frac{h_m}{p_n^{i_n}} + h = \frac{\bar{h}_2}{p_2^{j_2}} + \cdots + \frac{\bar{h}_n}{p_n^{j_n}} + h,$$

et nous pouvons alors achever la démonstration par récurrence.

L'expression donnée au théorème 8, est appelée *décomposition en éléments simples de R.*

Les polynômes irréductibles $p_1, \dots, p_n$ du théorème 8 peuvent être connus d'une façon en quelque sorte plus précise; le théorème suivant fournit des informations supplémentaires les concernant, et concernant également h.

Théorème 9. *Avec les notions du théorème 8, considérons une fraction rationnelle R exprimée sous la forme $R = g/f$, où f et g sont des polynômes étrangers, et $f \neq 0$.*
Supposons que tous les entiers $i_1, \dots, i_n$ soient > 0. Alors

$$f = p_1^{i_1} \cdots p_n^{i_n}$$

est la décomposition de f en éléments irréductibles. De plus, si $\deg g < \deg f$, $h = 0$.

Démonstration. Si on réduit la décomposition de R en éléments du théorème au même dénominateur, nous obtenons

$$(*) \qquad R = \frac{h_1 p_2^{i_2} \cdots p_n^{i_n} + \cdots + h_n p_1^{i_1} \cdots p_{n-1}^{i_{n-1}} + h p_1^{i_1} \cdots p_n^{i_n}}{p_1^{i_1} \cdots p_n^{i_n}}$$

Le polynôme p_ν ne divise donc pas le numérateur de droite dans (*), pour tout indice $\nu = 1, \dots, n$. En effet, p_ν divise chaque terme de ce numérateur *sauf* le terme

$$h_\nu p_1^{i_1} \cdots \widehat{p_\nu^{i_\nu}} \cdots p_n^{i_n}$$

(où le chapeau sur $p_\nu^{i_\nu}$ indique l'omission de ce facteur). Cela vient de l'hypothèse que p_ν ne divise pas h_ν. Le numérateur et le dénominateur de droite dans (*) sont donc premiers entre eux, prouvant ainsi notre première assertion.

Pour ce qui est de la seconde assertion, en posant encore $R = g/f$, nous avons $f = p_1^{i_1} \cdots p_n^{i_n}$ et

$$g = Rf = h_1 p_2^{i_2} \cdots p_n^{i_n} + \cdots + h_n p_1^{i_1} \cdots p_{n-1}^{i_{n-1}} + h p_1^{i_1} \cdots p_n^{i_n}$$

Supposons que $\deg g < \deg f$. Chaque terme de la somme précédente est alors de degré inférieur à degré de f, sauf peut-être le dernier terme

$$hf = hp_1^{i_1} \cdots p_n^{i_n}$$

Si $h \neq 0$, alors ce dernier terme est de degré $\geqq \deg f$ et nous obtenons maintenant

$$hf = g - h_1 p_2^{i_2} \cdots p_n^{i_n} - \cdots - h_n p_1^{i_1} \cdots p_{n-1}^{i_{n-1}}$$

où le membre de gauche est de degré $\geqq \deg f$ et le membre de droite de degré $< \deg f$. C'est impossible. Donc $h = 0$, comme il fallait le montrer.

Remarque. Étant donnée une fraction rationnelle $R = g/f$, où f et g sont des polynômes étrangers, on peut écrire en utilisant l'algorithme d'Euclide

$$g = g_1 f + g_2,$$

où g_1 et g_2 sont des polynômes et où $\deg g_2 < \deg f$. On a alors

$$\frac{g}{f} = \frac{g_2}{f} + g_1,$$

et on peut appliquer le théorème 9 à la fraction rationnelle g_2/f. Il est toujours utile lorsqu'on étudie des fractions rationnelles d'effectuer dès le début cette division pour ramener cette étude au cas où le degré du numérateur est plus petit que celui du dénominateur.

Exemple 1. Soient $\alpha_1, \ldots, \alpha_n$ des éléments distincts de K. Il existe des éléments $a_1, \ldots, a_n \in K$ tels que

$$\frac{1}{(t-\alpha_1)\cdots(t-\alpha_n)} = \frac{a_1}{t-\alpha_1} + \cdots + \frac{a_n}{t-\alpha_n}.$$

En effet, on peut, dans le cas présent, appliquer les théorèmes 8 et 9, avec $g = 1$ et en tirer que $\deg g < \deg f$. A l'exercice 14 du paragraphe précédent, nous avons montré comment déterminer a_i, par une méthode particulière.

On peut décomposer les expression $h_\nu/p_\nu^{i_\nu}$ de la décomposition en éléments simples, de la façon particulière que nous allons maintenant exposer.

Théorème 10. *Soit φ un polynôme non nul à coefficients dans K. Soit h un polynôme quelconque à coefficients dans K. Il existe des polynômes $\psi_0, \ldots, \psi_m$ tels que*

$$h = \psi_0 + \psi_1 \varphi + \cdots + \psi_m \varphi^m,$$

avec $\deg \psi_i < \deg \varphi$, *pour tout* $i = 0, \ldots, m$. *Les polynômes* $\psi_0, \ldots, \psi_m$ *sont déterminés de façon unique par ces conditions.*

Démonstration. Démontrons l'existence de $\psi_0, \ldots, \psi_m$ par récurrence sur le degré de h. On peut écrire, par division euclidienne,

$$h = q\varphi + \psi_0$$

pour des polynômes q et ψ_0, avec $\deg \psi_0 < \deg \varphi$. On a alors $\deg q < \deg h$, de telle sorte qu'on peut écrire, par hypothèse de récurrence

$$q = \psi_1 + \psi_2\varphi + \cdots + \psi_m\varphi^{m-1}$$

pour des polynômes ψ_i tels que $\deg \psi_i < \deg \varphi$. On obtient par substitution

$$h = (\psi_1 + \psi_2\varphi + \cdots + \psi_m\varphi^{m-1})\varphi + \psi_0$$

qui conduit à l'expression souhaitée.

Pour ce qui est de l'unicité, remarquons d'abord que, dans l'expression donnée dans le théorème, à savoir

$$h = \psi_0 + \psi_1\varphi + \cdots + \psi_m\varphi^m = \psi_0 + \varphi(\psi_1 + \cdots + \psi_m\varphi^{m-1})$$

le polynôme ψ_0 est nécessairement le reste de la division de h par φ, de sorte que l'unicité est assurée par la division euclidienne. En écrivant $h = q\varphi + \psi_0$, on conclut alors que

$$q = \psi_1 + \cdots + \psi_m\varphi^{m-1}$$

et q est déterminé de façon unique. Les polynômes $\psi_1, \ldots, \psi_m$ sont donc déterminés de façon unique par hypothèse de récurrence, et la démonstration est achevée.

L'expression de h en fonction des puissances de φ comme on l'a donné au théorème 10 s'appelle son *développement φ-adique.* On peut appliquer cela au cas où φ est un polynôme irréductible p, auquel cas son expression est le développement p-adique de h. Supposons que

$$h = \psi_0 + \psi_1 p + \cdots + \psi_m p^m$$

soit son développement p-adique. En divisant par p^i pour un certain entier $i > 0$, on obtient le théorème qui suit

Théorème 11. *Soient h un polynôme et p un polynôme irréductible à coefficients dans le corps K. Soit i un entier > 0. Il existe une expression unique*

$$\frac{h}{p^i} = \frac{g_{-i}}{p^i} + \frac{g_{-i+1}}{p^{i-1}} + \cdots + g_0 + g_1 p + \cdots + g_s p^s$$

où les g_μ sont des polynômes de degré $< \deg p$.

Nous avons adopté, dans le théorème 11, un numérotage de $g_{-i}, g_{-i+1}, \ldots$, de façon à l'ajuster à l'exposant de p intervenant au dénominateur. A part cela et modulo l'indexation, ces polynômes g ne sont rien d'autre que les $\psi_0, \psi_1, \ldots$ trouvés dans le développement p-adique de h.

Corollaire *Soit $\alpha \in K$, et soit h un polynôme à coefficients dans K. On a alors*

$$\frac{h(t)}{(t+\alpha_i)^i} = \frac{a_{-i}}{(t-\alpha)^i} + \frac{a_{-i+1}}{(t-\alpha)^{i-1}} + \cdots + a_0 + a_1(t-\alpha) + \cdots,$$

où les a_μ sont des éléments de K, déterminés de façon unique.

Démonstration. Dans ces conditions $p(t) = t - \alpha$ est de degré 1, de sorte que les coefficients du développement p-adique de h doivent être des constantes.

Exemple 2. On peut, pour déterminer la décomposition en éléments simples d'une fraction rationnelle donnée, résoudre un système d'équations linéaires. Donnons un exemple. On veut déterminer les constantes a et b telles que

$$\frac{1}{(t-1)(t-2)} = \frac{a}{t-1} + \frac{b}{t-2}.$$

En réduisant le second membre au même dénominateur, on a

$$\frac{1}{(t-1)(t-2)} = \frac{a(t-2)+b(t-1)}{(t-1)(t-2)}.$$

En égalant les numérateurs, on doit avoir

$$\begin{aligned} a + b &= 0, \\ -2a - b &= 1. \end{aligned}$$

On peut alors résoudre en a et b pour obtenir $a = -1$ et $b = 1$. On traite de façon similaire le cas général.

Exercices

1. Déterminez les décompositions en éléments simples des fractions rationnelles suivantes:

(a) $\dfrac{t+1}{(t-1)(t-2)}$; (b) $\dfrac{1}{(t+1)(t^2+2)}$.

2. Soit $R = g/f$ une fraction rationnelle telle que $\deg g < \deg f$. Soit

$$\frac{g}{f} = \frac{h_1}{p_1^{i_1}} + \cdots + \frac{h^n}{p_n^{i_n}}$$

sa décomposition en éléments simples. Soit $d_\nu = \deg p_\nu$. Montrez que les coefficients de $h_1, \ldots, h_n$ sont solutions d'un système d'équations linéaires, tel que le nombre de variables soit égal au nombre d'équations, i.e. tel que

$$\deg f = i_1 d_1 + \cdots + i_n d_n.$$

Le théorème 9 montre que ce système a une solution unique.

3. Trouvez le développement $(t-2)$-adique des polynômes suivants:

(a) $t^2 - 1$; (b) $t^3 + t - 1$; (c) $t^2 + 3$; (d) $t^4 + 2t^3 - t + 5$.

4. Trouvez le développement $(t-3)$-adique des polynômes de l'exercice 3.

§5. *Polynômes à coefficients entiers*

Les polynômes à coefficients dans **Z** forment un anneau particulièrement intéressant. Nous allons démontrer quelques propriétés particulières de tels polynômes, propriétés qui vont conduire à un important critère d'irréductibilité concernant les polynômes à coefficients rationnels. Si A est un anneau, contenu dans un corps, on désigne par $A[t]$ l'ensemble des polynômes à coefficients dans A. Il s'agit, de façon claire, d'un anneau, appelé anneau des polynômes à coefficients dans A.

Lemme 1. *Soit f un polynôme non nul sur le corps des rationnels. Il existe alors un nombre rationnel $r \neq 0$ tel que rf possède des coefficients entiers, qui sont premiers entre eux.*

Démonstration. Écrivons

$$f(t) = a_n t^n + \cdots + a_0,$$

où $a_0, \ldots, a_n$ sont des nombres rationnels, et où $a_n \neq 0$. Soit d un dénominateur commun de $a_0, \ldots, a_n$. Le polynôme df possède donc des coefficients entiers, à savoir $da_n, \ldots, da_0$. Soit b le plus grand commun diviseur de $da_n, \ldots, da_0$. Le polynôme

$$\frac{d}{b} f(t) = \frac{da_n}{b} t^n + \cdots + \frac{da_0}{b}$$

possède alors des coefficients entiers premiers entre eux, comme il fallait le montrer

Lemme 2. *Soient f et g deux polynômes non nuls à coefficients dans* **Z**; *supposons que f a des coefficients premiers entre eux, ainsi que g. Alors fg a aussi des coefficients premiers entre eux.*

Démonstration. Écrivons

$$f(t) = a_n t^n + \cdots + a_0, \quad a_n \neq 0,$$
$$g(t) = b_m t^m + \cdots + b_0, \quad b_m \neq 0,$$

où les a_i sont étrangers, ainsi que les b_i; soit p un nombre premier. Il va suffire de démontrer que p divise certains coefficients de fg. Soit r le plus grand entier tel que $0 \leqq r \leqq n$ avec $a_r \neq 0$, et tel que p ne divise pas a_r. De la même façon, soit b_s le coefficient de g le plus à gauche, non nul, tel que p ne divise pas b_s. Considérons le coefficient de t^{r+s} dans $f(t)g(t)$. Il est égal à

$$c = a_r b_s + a_{r+1} b_{s-1} + \cdots$$
$$+ a_{r-1} b_{s+1} + \cdots$$

et p ne divise pas $a_r b_s$. Cependant, p divise tout autre terme non nul de la somme puisque chacun de ces termes est de la forme

$$a_j b_{r+s-i}$$

avec a_i à gauche de a_r (i.e. $i > r$), ou de la forme

$$a_{r+s-j}b_j$$

avec $j > s$ (i.e. dans le cas où b_j est à gauche de b_s). Par suite, p ne divise pas c, et notre lemme est démontré.

***Théorème* 12.** (Gauss). *Soit f un polynôme à coefficients premiers entre eux de* deg $\geqq 1$. *Si f est réductible sur Q c'est-à-dire si l'on peut écrire $f = gh$ avec g, $h \in \mathbf{Q}[t]$, et* $\deg g \geqq 1, \deg h \geqq 1$, *alors il existe des nombres rationnels r et s tels que, si l'on pose $g_1 = rg$ et $h_1 = sh$, alors g_1 et h_1 sont à coefficients entiers, vérifiant $f = g_1h_1$.*

Démonstration. D'après le lemme 1, r et s sont des nombres rationnels non nuls tels que rg et sh sont à coefficients entiers, étrangers. Soient $g_1 = rg$ et $h_1 = sh$. Alors

$$f = \frac{1}{r}g_1 \frac{1}{s}h_1,$$

et par suite, $rsf = g_1h_1$. D'après le lemme 2, g_1h_1 possède des coefficients premiers entre eux. Puisqu'on suppose que les coefficients de f sont étrangers, il s'ensuit immédiatement que rs lui-même doit être un entier, et ne peut être divisible par aucun nombre premier. De là vient que $rs = \pm 1$, et qu'en divisant par exemple g_1 par rs, on obtient le résultat

***Théorème* 13.** (Critère d'Eisenstein.) *Soit*

$$f(t) = a_nt^n + \cdots + a_0$$

un polynôme de degré $n \geqq 1$, à coefficients entiers. Soit p un nombre premier. Supposons que

$$a_n \not\equiv 0 \pmod p, \qquad a_i \equiv 0 \pmod p \text{ pour tout } i < n$$
$$a_0 \not\equiv 0 \pmod{p^2})$$

Le polynôme f est alors irréductible sur **Q**.

Démonstration. Divisons d'abord f par le plus grand commun diviseur de ses coefficients et supposons maintenant que f a des coefficients premiers entre eux. D'après le théorème 5, on doit montrer que f ne peut pas s'exprimer comme produit gh de polynômes g et h à coefficients entiers, tels que deg g et deg h soient $\geqq 1$. Supposons qu'on peut le faire, et écrivons

$$g(t) = b_dt^d + \cdots + b_0,$$
$$h(t) = c_mt^m + \cdots + c_0,$$

où d, $m \geqq 1$ et $b_dc_m \neq 0$. Puisque $b_0c_0 = a_0$ est divisible par p mais pas par p^2, il s'ensuit que l'un des nombres b_0 ou c_0 n'est pas divisible par p, par exemple, b_0. Alors $p|c_0$. Puisque $c_mb_d = a_n$ n'est pas divisible par p, il s'ensuit que p ne divise

pas c_m. Soit c_r le coefficient de h le plus éloigné vers la droite tel que $c_r \not\equiv 0 \pmod{p}$. Alors $r \neq n$ et

$$a_r = b_0 c_r + b_1 c_{r-1} + \cdots$$

Puisque $p \nmid b_0 c_r$ mais puisque p divise tout autre terme de la somme, on en conclut que $p \nmid a_r$, contradiction qui démontre notre théorème.

Exemple. Le polynôme $t^5 - 2$ est irréductible sur le corps des rationnels, par application immédiate du théorème 13.

Remarque. Les démonstrations qui prouvent l'existence et l'unicité des décompositions dans les anneaux de polynômes et dans **Z** (donnée au chapitre I pour cette dernière) sont très semblables. Nous avons d'abord utilisé l'algorithme d'Euclide dans les deux cas, pour montrer que tout idéal peut être engendré par un seul élément. Il y a très peu d'exemples d'anneaux dans lesquels soit valable un tel algorithme d'Euclide. Cependant, les raisonnements qui ont suivi ne dépendaient que de ce qu'entraînait cet algorithme sur l'engendrement des idéaux. Il est donc utile de donner un nom à de tels anneaux: un anneau A est dit *anneau principal* s'il est intègre, et si tout idéal peut être engendré par un seul élément. On peut prouver le théorème de l'existence et de l'unicité des décompositions dans les anneaux principaux, et aussi démontrer les analogues des théorèmes 12 et 13 sur de tels anneaux. Il existe beaucoup d'exemples de tels anneaux, et cela vaut donc la peine d'expliciter de telles axiomatisations. Nous renvoyons le lecteur à des textes d'un niveau plus élevé, où il trouvera un exposé systématique de ces questions.

Exercices

1. Soit $f(t) = t^n + \cdots + a_0$ un polynôme de degré $n \geqq 1$, à coefficients entiers, de coefficient dominant égal à 1. Montrez que si f a une racine rationnelle, alors cette racine est, en fait, un entier, et que cet entier divise a_0.

2. Déterminez lesquels parmi les polynômes suivants sont irréductibles sur le corps des nombres rationnels:

(a) $t^3 - t + 1$, (b) $t^3 + 2t + 10$, (c) $t^3 - t - 1$, (d) $t^3 - 2t^2 + t + 15$.

3. Déterminez lesquels parmi les polynômes suivants sont irréductibles sur le corps des rationnels:

(a) $t^4 + 2$, (b) $t^4 - 2$, (c) $t^4 + 4$, (d) $t^4 - t + 1$.

4. Soit $f(t) = a_n t^n + \cdots + a_0$ un polynôme de degré $n \geqq 1$ à coefficients entiers, supposés premiers entre eux, avec $a_0 \neq 0$. Si b/c est un nombre rationnel exprimé comme quotient d'entiers premiers b et c non nuls, et si $f(b/c) = 0$, montrez que c divise a_n et que b divise a_0. (Ce résultat nous permet de déterminer toutes les racines rationnelles possibles de f, puisqu'il n'y a qu'un nombre fini de diviseurs de a_0 et de a_n.)

5. Déterminez toutes les racines rationnelles des polynômes suivants:

(a) $t^7 - 1$, (b) $t^8 - 1$, (c) $2t^2 - 3t + 4$, (d) $3t^3 + t - 5$, (e) $2t^4 - 4t + 3$.

§6. *Éléments transcendants*

Soit K un corps; supposons que K est un sous-anneau d'un anneau commutatif E. Soit $x \in E$. Utilisons le symbole $K[x]$ pour désigner l'ensemble de tous les éléments

$$a_0 + a_1 x + \cdots + a_n x^n,$$

où tous les $a_i \in K$ et tous les entiers n sont $\geqq 0$. Il est alors clair que $K[x]$ est un anneau, qu'on dit *engendré* par x sur K. Soit

$$f(t) = a_0 + a_1 t + \cdots + a_n t^n$$

un polynôme à coefficients dans K. On *définit* alors $f(x)$ par

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n.$$

(Puisque les coefficients de f sont déterminés de manière unique, cette application est bien définie.) Nous prétendons donc maintenant que l'application

$$f \mapsto f(x)$$

est un homomorphisme d'anneaux de $K[t]$ dans E. Cela est facile à vérifier. Soit, par exemple,

$$g(t) = b_0 + b_1 + b_1 t + \cdots + b_m t^m$$

un autre polynôme à coefficients dans K. Alors

$$\begin{aligned} f(x)g(x) &= (a_0 + a_1 x + \cdots + a_n x^n)(b_0 + \cdots + b_m x^m) \\ &= \sum_{i=0}^{n} \sum_{j=0}^{m} a_i x^i b_j x^j = \sum_{i=0}^{m} \sum_{j=0}^{m} a_i b_j x^{i+j} \\ &= \sum_{k=0}^{m+n} \Big(\sum_{i+j=k} a_i b_j \Big) x^k \\ &= (fg)(x) \end{aligned}$$

Cela montre que $(fg)(x) = f(x)g(x)$. On constate, même plus facilement, que $(f + g)(x) = f(x) + g(x)$, et que l'application envoie le polynôme constant 1 sur l'élément 1 de E. Ainsi nous avons un homomorphisme d'anneaux dont on dit qu'il est obtenu par *substitution de x dans les polynômes.*

Si cet homomorphisme d'anneaux est un isomorphisme, on dit que x est *transcendant sur K*, et on obtient un isomorphisme de $K[t]$ dans $K[x]$.

La fonction I de K dans lui-même telle que $I(t) = t$, est transcendante sur K et, quand nous écrivons un polynôme $f(t)$, le symbole t est essentiellement une simple notation de cette fonction I. La chose importante est ici que nous avons été capables de prouver l'existence d'éléments transcendants sur K en exhibant ces derniers comme fonctions polynomiales.

Si x et y sont transcendants sur K, alors $K[x]$ et $K[y]$ sont isomorphes, puisque tous deux isomorphes à l'anneau de polynômes $K[t]$. Notre définition des polynômes est donc simplement une façon concrète de traiter de l'anneau engendré sur K par un élément transcendant.

Quand K est un corps *fini*, on ne peut plus se servir des fonctions pour le faire. En effet, si K possède q éléments, l'application $t \mapsto t^q$ est l'application identique de K sur lui-même (parce que le groupe multiplicatif des éléments non nuls de K possède $q - 1$ éléments, de sorte que si $x \in K$, $x \neq 0$, alors $x^{q-1} = 1$ et $x^q = x$). Il est cependant possible de démontrer l'existence d'éléments transcendants, par différents expédients techniques. Nous allons maintenant en décrire un.

Partons d'un symbole X et désignons par X^n un nouveau symbole pour chaque entier $n \geqq 0$.

Nous voulons définir des sommes. Nous nous demandons d'abord: qu'exigeons-nous d'un polynôme

$$a_0X^0 + \cdots + a_nX^n?$$

Nous exigeons qu'il soit une somme (en un certain sens) de termes a_iX^i. Nous devons donc donner une définition de chaque terme a_iX^i. Si $a \in K$, on veut que aX^n soit complètement déterminé par a et n, et nous voyons qu'il n'y rien de plus qu'une correspondance de a avec X^n, et de 0 avec X^i pour tout $i \neq n$. Mais une correspondance n'est rien d'autre qu'une fonction. Cela nous indique comment définir aX^n.

On définit aX^n comme étant la fonction définie sur l'ensemble $\{X^0, X^1, ...\}$ qui associe a avec X^n et 0 avec tout X^i pour $i \neq n$. On appelle aX^n *monôme* à coefficients dans K.

Désignons par $K[X]$ l'ensemble de toutes les sommes de tels monômes

$$a_0X^0 + \cdots + a_nX^n$$

où $a_i \in K$, qui est l'ensemble de toutes les fonctions de $\{X^0, X^1, ...\}$ dans K qu'on peut écrire comme sommes finies de monômes. On appelle $K[X]$ l'ensemble des *polynômes formels* à coefficients dans K, et X une *variable* ou une *indéterminée sur* K. Si deux sommes

$$\Sigma\, a_iX^i \quad \text{et} \quad \Sigma\, b_iX^i$$

sont égales, alors nous avons, par définition, $a_i = b_i$ pour tout i. On appelle alors coefficients du polynôme $\Sigma\, a_iX^i$, les éléments $a_0, a_1, ...$ Les sommes étant définies, définissons maintenant le produit de deux polynômes d'une façon naturelle, en posant que

$$(a_0X^0 + \cdots + a_nX^n)(b_0X^0 + \cdots + b_mX^m)$$

est par définition, le polynôme

$$c_0X^0 + \cdots + c_{m+n}X^{m+n}$$

$$c_r = \sum_{i+j=r} a_ib_j = a_0b_r + a_1b_{r-1} + \cdots + a_rb_0.$$

C'est alors une question de routine que de démontrer que $K[X]$ est un anneau, c'est-à-dire que notre addition et notre multiplication satisfont les axiomes d'anneaux. Nous allons démontrer l'associativité de la multiplication, et laisser la démonstration des autres axiomes comme exercices.

Soient

$$f(X) = \Sigma\, a_i X^i,\ g(X) = \Sigma\, b_i X^i,\ h(X) = \Sigma d_i X^i$$

des polynômes. Alors

$$f(X)g(X) = \Sigma\, c_i X^i$$

avec

$$c_r = \sum_{i+j=r} a_i b_j,$$

et

$$\big(f(X)g(X)\big)\big(h(X)\big) = \Sigma\, e_i X^i$$

où, par définition,

$$e_s = \sum_{r+k=s} c_r d_k = \sum_{r+k=s}\left(\sum_{i+j=r} a_i b_j\right) d_k$$
$$= \sum_{i+j+k=s} a_i b_j d_k.$$

On prend la dernière somme sur tous les triplets (i, j, k) d'entiers $\geqq 0$, tels que $i + j + k = s$. Si on calcule maintenant $f(X)\big(g(X)h(X)\big)$ d'une façon analogue, on trouve exactement les mêmes coefficients que pour $\big(f(X)g(X)\big)h(X)$, prouvant par là l'associativité.

C'est également un exercice routinier de démontrer que l'application

$$a \mapsto aX^0$$

est un plongement de K dans l'anneau $K[X]$. Du point de vue notation, on écrit habituellement X au lieu de $1X$, de sorte que $X^n X^m = X^{n+m}$.

A partir d'ici, on peut montrer tous les théorèmes du §1 au §4 sans aucun changement dans les démonstrations, la seule exception étant la dernière assertion du théorème 1 : si f et g sont deux polynômes formels à coefficients dans un corps K fini, il n'est pas en général vrai que si $f(t) = g(t)$ pour tout $t \in K$, alors $f = g$. Par exemple, si K possède q éléments, alors les deux polynômes

$$X \quad \text{et} \quad X^q$$

conduisent à la même fonction de K dans lui-même, mais ne sont pas égaux en tant que polynômes formels. Le premier est de degré 2, et le second de degré q.

Le seul rôle joué par la dernière assertion du théorème 1, dans ce qui précède. est de montrer qu'une fonction polynomiale détermine de manière unique ses coefficients, et pour cela il nous a fallu un corps infini. Par la présente approche des polynômes, cela résulte immédiatement de la définition.

En mathématiques on veut habituellement dire par polynôme, polynôme formel, à moins d'autres précisions. Quand on veut traiter de fonctions polynomiales, on le dit explicitement.

Soit K un sous-corps, que l'on peut supposer fini, d'un corps E. Soit $x \in E$. L'ensemble de tous les éléments $f(x)$, où f est un polynôme formel à coefficients dans K et où $x \in E$ est évidemment un sous-anneau de E, désigné par $K[x]$. L'ensemble de tous les quotients

$$\frac{f(x)}{g(x)},$$

où f et g sont des polynômes à coefficients dans K tels que $g(x) \neq 0$, est évidemment un sous-corps de E que nous désignons par $K(x)$.

Soit X une indéterminée sur K, comme ci-dessus. L'application

$$f(X) \mapsto f(x)$$

est un homomorphisme de $K[X]$ sur l'anneau $K[x]$. Si cet homomorphisme est un isomorphisme, i.e. si x est transcendant sur K, alors on peut prolonger cette application au corps des fractions de $K[X]$, qui est désigné par $K(X)$, et, dans ces conditions, $K(X)$ et $K(x)$ sont isomorphes.

Au chapitre VI, nous étudions le cas où l'application

$$f(X) \mapsto f(x)$$

n'est *pas* un isomorphisme.

EXERCICES

1. Soit K un corps fini à q éléments. Si f et g sont des polynômes sur K de degré $< q$, et si $f(x) = g(x)$, pour tout $x \in K$, montrez que $f = g$ (en tant que polynômes formels).

2. Soit K un corps fini à q éléments. Soit f un polynôme à coefficients dans K. Montrez qu'il existe un polynôme f^* à coefficients dans K de degré $< q$ tel que

$$f^*(x) = f(x),$$

pour tout $x \in K$;

3. Soit K un corps fini à q éléments. Soit $a \in K$. Montrez qu'il existe un polynôme f à coefficients dans K tel que $f(a) = 0$ et $f(x) = 1$ pour un x de K différent de a. [*Indication:* $(X - a)^{q-1}$.]

4. Soit K un corps fini à q éléments. Soit $a \in K$. Montrez qu'il existe un polynôme f à coefficients dans K tel que $f(a) = 1$ et $f(x) = 0$ pour un x de K différent de a.

5. Soit K un corps fini à q éléments. Soit $\varphi : K \to K$ une fonction quelconque de K dans lui-même. Montrez qu'il existe un polynôme f sur K tel que $\varphi(x) = f(x)$, pour tout $x \in K$.

$f(x) = \sum_{i=1}^{q}$ [illegible] $[1-(x-k_i)^{q-1}]$

§7. *Polynômes à plusieurs variables*

Soit n un entier $\geqq 1$. Soit K un corps (supposé infini selon nos conventions). Nous définissons un *polynôme à n variables sur K* comme étant une fonction

$$f: K^n \to K$$

qu'on peut écrire sous la forme

$$f(t_1, \ldots, t_n) = \text{une somme finie de termes du type } a_{i_1\cdots i_n} t_1^{i_1} \cdots t_n^{i_n},$$

où les $a_{i_1\cdots i_n} \in K$, et où les $i_1, \ldots, i_n$ sont des entiers $\geqq 0$. Nous abrégeons une telle somme par

$$f(t_1, \ldots, t_n) = \sum_{(i)} a_{(i)} M_{(i)}(t_1, \ldots, t_n)$$

où

$$M_{(i)}(t_1, \ldots, t_n) = t_1^{i_1} \cdots t_n^{i_n}.$$

Si nous regroupons tous les termes ayant le même exposant i_n en t_n, on peut écrire f sous la forme

$$f(t_1, \ldots, t_n) = \sum_{i_n = 0}^{d_n} \left(\sum_{i_1, \cdots, i_{n-1}} a_{(i)} t_1^{i_1} \cdots t_{n-1}^{i_{n-1}} \right) t_n^{i_n}$$

et nous voyons alors qu'on peut écrire

$$f(t_1, \ldots, t_n) = \sum_{j=0} f_j(t_1, \ldots, t_{n-1}) t_n^j,$$

où les f_j sont des polynômes à $n - 1$ variables.

***Théorème* 14.** *Soit $f(t_1, \ldots, t_n)$ un polynôme sur le corps K, comme ci-dessus. Soient $S_1, \ldots, S_n$ des sous-ensembles infinis de K; supposons que $f(t_1, \ldots, t_n) = 0$ pour tous les $t_i \in S_i$. On a alors $a_{(i)} = 0$ pour tous les n-uples (i) d'entiers $\geqq 0$.*

Démonstration. Par récurrence sur n. Si $n = 1$, le théorème est un cas particulier du théorème 1, §1. Soit $n > 1$; supposons le théorème démontré au rang $n - 1$. Soient $t_1, \ldots, t_{n-1}$ des éléments de $S_1, \ldots, S_{n-1}$ respectivement. On obtient alors un polynôme en une seule variable

$$g(t) = \sum_{j=0}^{d_n} b_j t^j,$$

où $b_j = f_j(t_1, \ldots, t_{n-1})$. De plus $f(t) = 0$, pour tout $t \in S_n$. Donc $b_j = 0$, pour tout j. Par hypothèse de récurrence, il en résulte que $a_{(i)} = 0$ pour tout (i), comme souhaité.

Du théorème 14, on tire que si f s'exprime comme

$$\sum_{(i)} a_{(i)} M_{(i)}(t_1, \ldots, t_n) = \sum_{(i)} b_{(i)} M_{(i)}(t_1, \ldots, t_n),$$

alors $a_{(i)} = b_{(i)}$ pour tout (i). (Retranchez et appliquez le théorème 14.) Les éléments $a_{(i)}$ de K sont déterminés de façon unique par f, et sont encore appelés *coefficients* de f.

Il sera commode d'abréger $f(t_1, \ldots, t_n)$ en $f(t)$. Il est clair que le produit de deux polynômes à plusieurs variables en est encore un, ainsi que la somme. Par suite,

les polynômes à n variables forment un sous-anneau de l'anneau des fonctions de K^n dans K, désigné par $K[t_1, ..., t_n]$, qu'on abrège aussi parfois en $K[t]$.

Exactement comme pour les polynômes à une seule variable, supposons que K soit un sous-anneau d'un anneau commutatif E. Soient $x_1, ..., x_n$ des éléments de E. Soit f un polynôme à n variables sur K, écrit comme précédemment, à coefficients $a_{(i)}$ et posons

$$f(x_1, ..., x_n) = \sum_{(i)} a_{(i)} x_1^{i_1} \cdots x_n^{i_n}.$$

L'application

$$f \mapsto f(x_1, ..., x_n)$$

est alors un homomorphisme d'anneaux de $K[t_1, ..., t_n]$ dans E. On note $K[x_1, ... x_n]$ son image. Si notre application est un isomorphisme de $K[t_1, ..., t_n]$ sur son image, on dit alors que $x_1, ..., x_n$ sont *algébriquement indépendants sur* K. On peut faire, pour plusieurs variables, les mêmes commentaires que ceux que nous avons faits dans le précédent paragraphe concernant la construction des polynômes sur des corps qui ne sont pas infinis.

De la même façon que dans le cas d'une seule variable, on peut définir $K(x_1, ..., x_n)$ comme ensemble de tous les quotients

$$\frac{f(x_1, ..., x_n)}{g(x_1, ..., x_n)}$$

où f et g sont des polynômes à plusieurs variables sur K, et où

$$g(x_1, ..., x_n) \neq 0$$

Il est clair que $K(x_1, ..., x_n)$ est un sous-groupe de E. (Donnez-en la démonstration en détail.)

***Théorème* 15.** *Soit K un sous-corps d'un corps E, et soit $x_1, ..., x_n \in E$. Soit $K_r = K(x_1, ..., x_r)$ pour $1 \leqq r \leqq n$. Si x_{r+1} est transcendant sur K_r pour chaque $r = 1, ..., n-1$, alors les éléments $x_1, ..., x_r$ sont algébriquement indépendants sur K, et réciproquement.*

Démonstration. Nous la laissons en exercice au lecteur.

CHAPITRE V

Espaces vectoriels et modules

§1. *Espaces vectoriels et bases*

Soit K un corps. *Un espace vectoriel* V *sur le corps* K est un groupe additif (abélien), sur lequel est définie une multiplication des éléments de V par ceux de K (les *scalaires* - N.d.T.) i.e. une application.

$$(x, v) \mapsto xv$$

de $K \times V$ dans V, satisfaisant aux conditions suivantes:

EV 1. *Si* 1 *est l'élément unité de* K, *alors* $1v = v$, *pour tout* $v \in V$.
EV 2. *Si* $c \in K$ *et si* $v, w \in V$, *alors* $c(v + w) = cv + cw$.
EV 3. *Si* $x, y \in K$ *et* $v \in V$, *alors* $(x + y)v = xv + yv$.
EV 4. $x, y \in K$ *et* $v \in V$, *alors* $(xy)v = x(yv)$.

Exemple 1. Soit V l'ensemble des fonctions continues à valeurs réelles définies sur l'intervalle $[0, 1]$. Cet ensemble V est un espace vectoriel sur $\mathbf{R}$. On définit comme d'habitude l'addition des fonctions: si f et g sont deux fonctions

$$(f + g)(t) = f(t) + g(t).$$

Pour tout c de $\mathbf{R}$, on pose $(cf)(t) = cf(t)$. C'est alors une simple question de routine de vérifier que nos quatre conditions EV 1, EV 2, EV 3 et EV 4 sont satisfaites.

Exemple 2. Soit E un ensemble non vide, et V l'ensemble de toutes les applications de V dans K. Cet ensemble V est un espace vectoriel sur K pour une addition des applications et une multiplication des applications par les éléments de V définies comme pour les fonctions de l'exemple précédent.

Exemple 3. Désignons par K^n le produit $K \times \cdots \times K$, i.e. l'ensemble des n-uples d'éléments de K. (Si $K = \mathbf{R}$, cet ensemble est l'espace euclidien habituel.) Définissons l'addition des n-uples composante par composante, ce qui veut dire que si

$$X = (x_1, \ldots, x_n) \quad \text{et} \quad Y = (y_1, \ldots, y_n)$$

sont des éléments de K^n, où $x_i, y_i \in K$, alors on pose

$$X + Y = (x_1 + y_1, \ldots, x_n + y_n).$$

Pour tout c de K, on pose

$$cX = (cx_1, ..., cx_n).$$

C'est encore une question de routine de vérifier que nos quatre conditions EV sont satisfaites par ces opérations.

Exemple 4. En faisant $n = 1$ dans l'exemple 3, on voit que K est un espace vectoriel sur lui-même.

Soit V un espace vectoriel sur un corps K. Soit $v \in V$. On a alors $0v = 0$. *Démonstration:* $0v + v = 0v + 1v = (0 + 1)v = 1v = v$. D'où, en ajoutant $-v$ aux deux membres, le fait que $0v = 0$.

Si $c \in K$ et $cv = 0$, alors que $c \neq 0$, alors $v = 0$. Pour le voir, on multiplie par c^{-1} pour obtenir $c^{-1}cv = 0$ et, par suite, $v = 0$.

On a $(-1)v = -v$. *Démonstration:*

$$(-1)v + v = (-1)v + 1v = (-1 + 1)v = 0v = 0.$$

D'où $(-1)v = -v$.

Soit V un espace vectoriel, et W un sous-ensemble de V. Nous disons que W est un *sous-espace vectoriel* ou un *sous-espace de V* si W est un sous-groupe (du groupe additif de V), et si, étant donnés $c \in K$ et $v \in W$, cv est encore élément de W. En d'autres termes, un sous-espace W de V est un sous-ensemble de V satisfaisant aux conditions suivantes:

(i) si v, w sont éléments de W, leur somme est un élément de V.
(ii) l'élément 0 de V est aussi élément de W.
(iii) si $v \in W$ et si $c \in K$, alors $cv \in W$.

Le sous-ensemble W est donc lui-même un espace vectoriel. En effet, les propriétés EV 1 à EV 4, étant satisfaites par tous les éléments de V sont *a fortiori* satisfaites par les éléments de W.

Soient V un espace vectoriel, et $w_1, ..., w_n$ des éléments de V. Soit W l'ensemble de tous les éléments

$$x_1w_1 + \cdots + x_nw_n$$

où $x_i \in K$. Alors W est un sous-espace de V, comme on le vérifie sans difficulté. Il est appelé *sous-espace engendré* par $w_1, ..., w_n$ et nous disons que $w_1, ..., w_n$ sont des *générateurs* de ce sous-espace.

Soit V un espace vectoriel sur le corps K et soient $v_1, ..., v_n$ des éléments de V. Nous dirons que $v_1, ..., v_n$ sont *linéairement dépendants* sur K s'il existe des éléments $a_1, ..., a_n$ dans K, non tous nuls, tels que

$$a_1v_1 + \cdots + a_nv_n = 0.$$

S'il n'existe pas de tels éléments, alors on dit que $v_1, ..., v_n$ sont linéairement indépendants sur K. Nous omettons souvent les mots «sur K».

Exemple 5. Soit $V = K^n$; considérons les vecteurs

$$v_1 = (1, 0, \ldots, 0)$$
$$\vdots$$
$$v_n = (0, 0, \ldots, 1);$$

ils sont linéairement indépendants. En effet, soient $a_1, \ldots, a_n$ des éléments de K tels que $a_1 v_1 + \cdots + a_n v_n = 0$. Puisque

$$a_1 v_1 + \cdots + a_n v_n = (a_1, \ldots, a_n),$$

il s'ensuit que tous les a_i sont nuls.

Exemple 6. Soit V l'espace vectoriel de toutes les fonctions d'une variable réelle t. Soient n fonctions $f_1(t), \ldots, f_n(t)$. Dire qu'elles sont linéairement dépendantes revient à dire qu'il existe n nombres réels $a_1, \ldots, a_n$, non tous nuls, tels que

$$a_1 f_1(t) + \cdots + a_n f_n(t) = 0$$

(pour toutes les valeurs de t).

Les deux fonctions e^t, e^{2t} sont linéairement indépendantes. Pour le montrer, nous allons supposer qu'il existe des nombres a et b tels que

$$ae^t + ba^{2t} = 0$$

(pour toutes les valeurs de t). Dérivons cette relation; il vient

$$ae^t + 2be^{2t} = 0.$$

Retranchons la première relation de la seconde. Nous obtenons $be^t = 0$, et donc $b = 0$. De la première relation s'ensuit alors que $ae^t = 0$, et donc que $a = 0$. Par suite, e^t et e^{2t} sont linéairement indépendantes.

Considérons de nouveau un espace vectoriel V quelconque sur un corps K. Soient $v_1, \ldots, v_n$ des éléments linéairement indépendants de V. Soient $x_1, \ldots, x_n$ et $y_1, \ldots, y_n$ des scalaires. Supposons que nous ayons

$$x_1 v_1 + \cdots + x_n v_n = y_1 v_1 + \cdots + y_n v_n.$$

Autrement dit, supposons que nous avons deux combinaisons linéaires des v_i égales. Nous devons alors avoir $x_i = y_i$ pour tout $i = 1, \ldots, n$. En effet, en soustrayant le membre de droite du membre de gauche, on obtient

$$x_1 v_1 - y_1 v_1 + \cdots + x_n v_n - y_n v_n = 0.$$

On peut encore écrire cette relation sous la forme

$$(x_1 - y_1)v_1 + \cdots + (x_n - y_n)v_n = 0.$$

Par définition, on dit alors avoir $x_i - y_i = 0$ pour tout $i = 1, \ldots, n$, prouvant par là ce que nous voulions montrer.

Nous appelons *base* de V sur K une suite $\{v_1, \ldots, v_n\}$ d'éléments de K qui engendrent V et sont linéairement indépendants.

Les vecteurs $v_1, \ldots, v_n$ de l'exemple 5 forment une base de K^n sur K.

Soit W l'espace vectoriel des fonctions engendré sur $\mathbf{R}$ par e^t et e^{2t}. L'ensemble e^t, e^{2t} est alors une base de W sur $\mathbf{R}$.

Soit V un espace vectoriel, et soit $\{v_1, ..., v_n\}$ une base de V. On peut représenter les éléments de V par les n-uples relatifs à cette base, de la façon qui suit. Si un élément v de V s'écrit en combinaison linéaire

$$v = x_1 v_1 + \cdots + x_n v_n$$

des éléments de la base, alors nous appelons *coordonnées* de v sur notre base les scalaires $(x_1, ..., x_n)$, et x_i est appelé la i-ième coordonnée. On dit que le n-uple $X = (x_1, ..., x_v)$ est le *vecteur des coordonnées* de v relativement à la base $\{v_1, ..., v_n\}$.

Soit, par exemple, V l'espace vectoriel des fonctions engendré par les deux fonctions e^t, e^{2t}. Les coordonnées de la fonction

$$3e^t + 5e^{2t}$$

sur la base $\{e^t, e^{2t}\}$ sont alors (3, 5).

Exemple 7. Montrons que les vecteurs (1, 1) et (−3, 2) sont linéairement indépendants sur $\mathbf{R}$.

Soient a et b deux nombres réels tels que

$$a(1, 1) + b(-3, 2) = 0.$$

En écrivant cette équation composante par composante, on trouve

$$\begin{aligned} a - 3b &= 0, \\ a + 2b &= 0. \end{aligned}$$

C'est là un système de deux équations que nous allons résoudre en a et b. En retranchant la seconde égalité de la première, on obtient $-5b = 0$, d'où $b = 0$. En substituant dans l'une ou l'autre équation, on trouve $a = 0$. Par suite, a et b sont tous deux nuls, et nos vecteurs linéairement indépendants.

Exemple 8. Trouvez les coordonnées de (1, 0) sur les vecteurs (1, 1) et (−1, 2). On doit trouver des nombres a et b tels que

$$a(1, 1) + b(-1, 2) = (1, 0).$$

En écrivant cette équation composante par composante, on trouve

$$\begin{aligned} a - b &= 1, \\ a + 2b &= 0. \end{aligned}$$

La résolution en a et b, de la manière habituelle, conduit à $b = -\frac{1}{3}$ et $a = \frac{2}{3}$. Par suite, les coordonnées de (1, 0) sur (1, 1) et (−1, 2) sont $(\frac{2}{3}, -\frac{1}{3})$.

Soit $\{v_1, ..., v_n\}$ un ensemble d'éléments d'un espace vectoriel V sur un corps K. Soit r un entier positif $\leqq n$. Nous dirons que $\{v_1, ..., v_r\}$ est un sous-ensemble *maximal* d'éléments linéairement indépendants de $\{v_1, ..., v_n\}$ si $v_1, ..., v_r$ sont linéairement indépendants, et si, de plus, étant donné un v_i quelconque pour $i > r$, les éléments $v_1, ..., v_r, v_i$ sont linéairement dépendants.

Le théorème suivant nous fournit un critère utile pour déterminer si un ensemble d'éléments d'un espace vectoriel est une base.

***Théorème* 1.** *Soit $\{v_1, \ldots, v_n\}$ un ensemble de générateurs d'un espace vectoriel V et $\{v_1, \ldots, v_r\}$ un sous-ensemble maximal constitué d'éléments linéairement indépendants. Alors $\{v_1, \ldots, v_r\}$ est une base de V.*

Démonstration. Nous devons démontrer que $v_1, \ldots, v_r$ engendrent V. Nous allons tout d'abord montrer que tous les v_i sont, pour $i > r$, des combinaisons linéaires de $v_1, \ldots, v_r$. Par hypothèse, étant donné un v_i, il existe $x_1, \ldots, x_r, y \in K$, non tous nuls, tels que

$$x_1 v_1 + \cdots + x_r v_r + y v_i = 0.$$

De plus, $y \neq 0$, parce que, sinon, nous aurions une relation de dépendance linéaire pour les $v_1, \ldots, v_r$. Par suite, on peut résoudre en v_i, pour obtenir

$$v_i = \frac{x_1}{-y} + \cdots + \frac{x_r}{-y} v_r,$$

montrant par là que v_i est bien combinaison linéaire de $v_1, \ldots, v_r$.

Soit maintenant v un élément quelconque de V. Il existe $c_1, \ldots, c_n \in K$ tels que

$$v = c_1 v_1 + \cdots + c_n v_n.$$

On peut, dans cette relation, remplacer chaque $v_i (i > r)$ par une combinaison linéaire de $v_1, \ldots, v_r$. Si nous le faisons, puis regroupons les termes, nous constatons que nous avons exprimé v comme combinaison linéaire de $v_1, \ldots, v_i$. Cela prouve que $v_1, \ldots, v_r$ engendrent V, et donc constituent une base de V.

Soient V et W deux espaces vectoriel sur K. Une application

$$f: V \to W,$$

est dite *application K-linéaire*, ou *homomorphisme d'espaces vectoriels*, si f satisfait les conditions suivantes:

Pour tout $x \in K$ et tous $v, v' \in V$,

$$f(v + v') = f(v) + f(v'), \quad f(xv) = xf(v).$$

L'application f est donc un homomorphisme de V dans W considéré comme groupes additifs, satisfaisant la condition supplémentaire $f(xv) = xf(v)$. Nous disons couramment «application linéaire» au lieu d'«application K-linéaire».

***Théorème* 2.** *Soient V et W des espaces vectoriels et $\{v_1, \ldots, v_n\}$ une base de V. Soient $w_1, \ldots, w_n$ des éléments de W. Il existe alors une unique application linéaire $f: V \to W$ telle que $f(v_i) = w_i$, pour tout i.*

Démonstration. L'application K-linéaire f est déterminée de façon unique, parce que, si

$$v = x_1 v_1 + \cdots + x_n v_n$$

est un élément de V, où les x_i appartiennent à K, nous devons alors nécessairement avoir

$$\begin{aligned} f(v) &= x_1 f(v_1) + \cdots + x_n f(v_n) \\ &= x_1 w_1 + \cdots + x_n w_n \end{aligned}$$

L'application f existe, car étant donné un élément v comme ci-dessus, on définit $f(v)$ comme étant $x_1 w_1 + \cdots + x_n w_n$. On doit alors vérifier que f est linéaire. Soit

$$w = y_1 v_1 + \cdots + y_n v_n$$

un élément de V, pour des $y_i \in K$. Alors

$$v + w = (x_1 + y_1)v_1 + \cdots + (x_n + y_n)v_n.$$

Par suite

$$\begin{aligned} f(v + w) &= (x_1 + y_1)w_1 + \cdots + (x_n + y_n)w_n \\ &= x_1 w_1 + y_1 w_1 + \cdots + x_n w_n + y_n w_n \\ &= f(v) + f(w) \end{aligned}$$

Si $c \in K$, alors $cv = cx_1 v_1 + \cdots + cx_n v_n$, et, par suite,

$$f(cv) = cx_1 w_1 + \cdots + cx_n w_n = cf(v).$$

Cela démontre que f est linéaire, et achève la démonstration du théorème.

Comme pour les groupes, on dit qu'une application linéaire $f : V \to W$ est un *isomorphisme* (i.e. un isomorphisme d'espaces vectoriels) s'il existe une application linéaire $g : W \to V$ telle que $g \circ f$ soit l'identité de V et $f \circ g$ celle de W. Le *noyau* d'une application linéaire est défini comme étant le noyau de l'application considérée comme homomorphisme de groupes. Démontrez enfin, à titre d'exercice, que le noyau et l'image d'une application linéaire sont des sous-espaces vectoriels.

EXERCICES

1. Montrez que les vecteurs suivants sont linéairement indépendants, sur **R** ou sur **C**.

(a) $(1, 1, 1)$ et $(0, 1, -1)$
(b) $(1, 0)$ et $(1, 1)$
(c) $(-1, 1, 0)$ et $(0, 1, 2)$
(d) $(2, -1)$ et $(1, 0)$
(e) $(\pi, 0)$ et $(0, 1)$
(f) $(1, 2)$ et $(1, 3)$
(g) $(1, 1, 0)$, $(1, 1, 1)$ et $(0, 1, -1)$
(h) $(0, 1, 1)$, $(0, 2, 1)$ et $(1, 5, 3)$

2. Exprimez les vecteurs X, donnés, comme combinaisons linéaires des vecteurs A et B, et trouvez les coordonnées de X sur A et sur B.

(a) $X = (1, 0)$, $A = (1, 1)$, $B = (0, 1)$
(b) $X = (2, 1)$, $A = (1, -1)$, $B = (1, 1)$
(c) $X = (1, 1)$, $A = (2, 1)$, $B = (-1, 0)$
(d) $X = (4, 3)$, $A = (2, 1)$, $B = (-1, 0)$

(Vous pouvez considérer les vecteurs ci-dessus comme éléments de $\mathbf{R}^2$ ou de $\mathbf{C}^2$. Les coordonnées sont les mêmes.)

3. Trouvez les coordonnées du vecteur X relativement aux vecteurs A, B, C.

(a) $X = (1,0,0)$, $A = (1,1,1)$, $B = (-1,1,0)$, $C = (1,0,-1)$
(b) $X = (1,1,1)$, $A = (0,1,-1)$, $B = (1,1,0)$, $C = (1,0,2)$,
(c) $X = (0,0,1)$, $A = (1,1,1)$, $B = (-1,1,0)$, $C = (1,0,-1)$.

4. Soient (a,b) et (c,d) deux vecteurs du plan. Si $ad - bc = 0$, montrez qu'ils sont linéairement indépendants.

5. Montrez que 1 et $\sqrt{2}$ sont linéairement indépendants sur le corps des rationnels.

6. Montrez que 1 et $\sqrt{3}$ sont linéairement indépendants sur le corps des rationnels.

7. Soit α un nombre complexe. Montrez que α est rationnel si et seulement si 1 et α sont linéairement dépendants sur le corps des rationnels.

8. Soient V et W deux espaces vectoriels sur le corps K; désignons par $\mathrm{Hom}_K(V, W)$ l'ensemble de toutes les applications linéaires de V dans W. Montrez que $\mathrm{Hom}_K(V, W)$ est un sous-groupe du groupe (additif) de toutes les applications de V dans W. Si $f : V \to W$ est une application linéaire, si $c \in K$, on définit cf par $(cf)(v) = cf(v)$. Montrez que $\mathrm{Hom}_K(V, W)$ est, dans ces conditions, un espace vectoriel sur K.

§2. *Dimension d'un espace vectoriel*

Le résultat principal de ce paragraphe assure que deux bases d'un espace vectoriel ont même nombre d'éléments. Pour le démontrer, nous établissons tout d'abord un résultat intermédiaire.

Théorème 3. *Soit V un espace vectoriel sur le corps K. Soit $\{v_1, ..., v_m\}$ une base de V sur K. Soient $w_1, ..., w_n$ des éléments de V : supposons que $n > m$. Alors, $w_1, ..., w_n$ sont linéairement indépendants.*

Démonstration. Supposons que $w_1, ..., w_n$ sont linéairement indépendants. Puisque $v_1, ..., v_m$ est une base, il existe des éléments $a_1, ..., a_m \in K$ tels que

$$w_1 = a_1 v_1 + \cdots + a_m v_m.$$

On sait, par hypothèse, que $w_1 \neq 0$, et par suite que l'un au moins des $a_i \neq 0$. Après une éventuelle re-indexation, on peut supposer, sans restreindre la généralité, que $a_1 \neq 0$. On peut alors résoudre en v_1, pour obtenir

$$a_1 v_1 = w_1 - a_2 v_2 - \cdots - a_m v_m,$$
$$v_1 = a_1^{-1} w_1 - a_1^{-1} a_2 v_2 - \cdots - a_1^{-1} a_m v_m.$$

Le sous-espace de V engendré par $w_1, v_2, ..., v_m$ contient v_1, et doit par conséquent être V en entier puisque $v_1, v_2, ..., v_m$ engendrent V. L'idée directrice est ici de poursuivre le procédé pas à pas et de remplacer successivement $v_2, v_3, ...,$ jusqu'à ce que tous les éléments $v_1, v_2, ..., v_m$ aient été pris, et qu'alors $w_1, ..., w_m$ engendrent V. Supposons donc par hypothèse de récurrence qu'il existe un entier r, avec $1 \leqq r < m$, tel qu'après ré-indexation convenable de $v_1, ..., v_m$, les éléments $w_1, ..., w_r, v_{r+1}, ..., v_m$ engendrent V. Il existe alors des éléments $b_1, ..., b_r, c_{r+1}, ..., c_m$ de K tels que

$$w_{r+1} = b_1 w_1 + \cdots + b_r w_r + c_{r+1} v_{r+1} + \cdots + c_m v_m.$$

Nous ne pouvons pas avoir $c_j = 0$ pour $j = r + 1, \ldots, m$ sinon nous pourrions en tirer une relation de dépendance entre $w_1, \ldots, w_{r+1}$ contredisant notre hypothèse. Après ré-indexation éventuelle de $v_{r+1}, \ldots, v_m$, on peut supposer, sans restreindre la généralité, que $c_{r+1} \neq 0$. On obtient alors

$$c_{r+1}v_{r+1} = w_{r+1} - b_1w_1 - \cdots - b_rw_r - c_{r+2}v_{r+2} - \cdots - c_mv_m.$$

En divisant par c_{r+1}, nous concluons que v_{r+1} est dans le sous-espace engendré par $w_1, \ldots, w_{r+1}, v_{r+2}, \ldots, v_m$. Il s'ensuit, par hypothèse de récurrence, que $w_1, \ldots, w_{r+1}, v_{r+2}, \ldots, v_m$ engendrent V. Nous avons ainsi démontré, par récurrence, que $w_1, \ldots, w_m$ engendrent V. Si on écrit

$$w_{m+1} = x_1w_1 + \cdots + x_mw_m,$$

où $x_i \in K$, on obtient une relation de dépendance linéaire

$$w_{m+1} - x_1w_1 - \cdots - x_mw_m = 0,$$

comme il fallait le démontrer.

***Théorème* 4.** *Soient V un espace vectoriel sur K, $\{v_1, \ldots, v_n\}$ et $\{w_1, \ldots, w_m\}$ deux bases de V. Alors $m = n$.*

Démonstration. D'après le théorème 3, on doit avoir $n \leqq m$ et $m \leqq n$, de telle sorte que $m = n$.

Si un espace vectoriel possède une base, alors toute autre base a le même nombre d'éléments. Ce nombre est appelé la *dimension* de V (sur K). Si V est l'espace vectoriel nul, on définit la dimension de V comme étant 0.

Corollaire. *Soit V un espace vectoriel de dimension n et soit W un sous-espace de V contenant n vecteurs linéairement indépendants. Alors $W = V$.*

Démonstration. Soit $v \in V$ et soient $w_1, \ldots, w_n$ des éléments linéairement indépendants de W. Les éléments $v, w_1, \ldots, w_n$ sont alors linéairement dépendants, de sorte qu'il existe $a, b_1, \ldots, b_n \in K$, non tous nuls, tels que

$$av + b_1w_1 + \cdots + b_nw_n = 0$$

On ne peut pas avoir $a = 0$, sinon $w_1, \ldots, w_n$ seraient linéairement dépendants. Alors

$$v = -a^{-1}b_1w_1 - \cdots - a^{-1}b_nw_n$$

est élément de W. Cela prouve que $V \subset W$, et qu'ainsi $V = W$.

§3. *Modules*

Considérons une généralisation de la notion d'espace vectoriel sur un corps K, à savoir celle de module sur un anneau. Soit un anneau A. Par *module* (à gauche) *sur A*, ou par *A-module*, on entend un groupe additif M donné avec une application

$A \times M \to M$, qui, à chaque couple (x, v) où $x \in A$ et $v \in M$, associe un élément xv de M, satisfaisant les quatre conditions:

MOD 1. *si e est l'élément unité de A, alors $ev = v$ pour tout $v \in M$.*
MOD 2. *si $x \in A$ et $v, w \in M$, alors $x(v + w) = xv + xw$.*
MOD 3. *si $x, y \in A$ et $v \in M$, alors $(x + y)v = xv + yv$.*
MOD 4. *si $x, y \in A$ et $v \in M$, alors $(xy)v = x(yv)$.*

Exemple 1. Tout idéal à gauche de A est un module. Le groupe additif réduit à 0 est un A-module pour tout anneau A.

Comme pour les espaces vectoriels, on a $0v = 0$ pour tout $v \in M$. (Remarquez que le 0 de $0v$ est l'élément nul de A, tandis que le 0 figurant dans l'autre membre de l'équation est l'élément nul du groupe additif M. Il n'y aura cependant aucune confusion résultant de l'utilisation du même symbole 0 pour les éléments nuls quels qu'ils soient.) Nous avons également $(-e)v = -v$, qu'on tire de la démonstration du même résultat pour les espaces vectoriels.

Soit M un A-module et N un sous-groupe de M. Nous disons que N est un *sous-module* de M si, lorsque $v \in N$ et $x \in A$, alors $xv \in N$. Il s'ensuit que N est alors lui-même un module.

Exemple 2. Soient M un module et $v_1, \dots, v_n$ des éléments de M. Soit N le sous-ensemble de M constitué de tous les éléments

$$x_1v_1 + \cdots + x_nv_n,$$

où $x_i \in A$. Alors N est un sous-module de M. En effet,

$$0 = 0v_1 + \cdots + 0v_n$$

de sorte que $0 \in N$. Si $y_1, \dots, y_n \in A$, alors

$$\begin{aligned} &x_1v_1 + \cdots + x_nv_n + y_1v_1 + \cdots + y_nv_n \\ &= (x_1 + y_1)v_1 + \cdots + (x_n + y_n)v_n \end{aligned}$$

est dans N. Enfin si $c \in A$, alors

$$c(x_1v_1 + \cdots + x_nv_n) = cx_1v_1 + \cdots + cx_nv_n$$

est dans N, et nous avons ainsi prouvé que N est un sous-module de M. On l'appelle le sous-module *engendré* par $v_1, \dots, v_n$ et nous disons que les éléments $v_1, \dots, v_n$ sont des *générateurs* de N.

Exemple 3. Soit M un groupe additif (abélien) et soit A un sous-anneau de End(M). (Nous avons défini End(M) au chapitre III, §1, comme l'anneau des homomorphismes de M dans lui-même.) Si on associe à chaque $f \in A$ et à chaque $v \in M$ l'élément $fv = f(v) \in M$, alors M est un A-module. La vérification des quatre conditions MOD 1 à MOD 4 se fait trivialement.

Réciproquement, étant donné un anneau A et un A-module M, on associe à chaque $x \in A$ l'application $\lambda_x : M \to M$ telle que $\lambda_x(v) = xv$, pour tout $v \in M$. La correspondance

$$x \mapsto \lambda_x$$

est un homomorphisme d'anneaux de A dans End(A), où End(A) est l'anneau des endomorphismes de M considéré comme groupe additif. Tout cela n'est autre qu'une nouvelle formulation des quatre conditions MOD 1 à MOD 4. Par exemple, on peut exprimer MOD 4 en écrivant

$$\lambda_{xy} = \lambda_x \lambda_y \quad \text{ou} \quad \lambda_{xy} = \lambda_x \circ \lambda_y,$$

puisque la multiplication dans End(M) est la composition des applications.

Attention. Il se peut que l'homomorphisme d'anneaux $x \mapsto \lambda_x$ ne soit pas injectif, de sorte qu'en général, en traitant du module M, on ne peut pas considérer A comme sous-anneau de End(M).

Soient A un anneau, M et M' des A-modules. On entend par application A-linéaire $f: M \to M'$ (ou *A-homomorphisme*) une application telle que pour tout $x \in A$ et pour tous v, $w \in M$, on ait

$$f(xv) = xf(v), \quad f(v + w) = f(v) + f(w).$$

Ainsi les applications A-linéaires généralisent les applications K-linéaires quand le module est un espace vectoriel sur un corps.

L'ensemble de toutes les applications A-linéaires de M dans M' sera désigné par $\mathrm{Hom}_A(M, M')$.

Exemple 4. Soient M, M', M'' trois A-modules. Si

$$f: M \to M' \quad \text{et} \quad g: M' \to M''$$

sont des applications A-linéaires, alors l'application composée $g \circ f$ est A-linéaire.

Par analogie avec les définitions précédentes, on dit qu'un A-homomorphisme $f: M \to M'$ est un *isomorphisme* s'il existe un A-homomorphisme $g: M' \to M$ tel que $g \circ f$ et $f \circ g$ sont les applications identiques de M et M' respectivement. Nous laissons au lecteur le soin de vérifier qu'*un A-homomorphisme qui est injectif et surjectif est un isomorphisme, et réciproquement.*

Comme pour les espaces vectoriels et les groupes additifs, nous avons très fréquemment à considérer l'ensemble des applications A-linéaires d'un module dans lui-même, et il est commode de donner un nom à ces applications. On les appelle des *A-endomorphismes* de M. L'ensemble des A-endomorphismes de M est noté $\mathrm{End}_A(M)$.

On supprime souvent le préfixe A- quand l'anneau de référence est clairement déterminé par le contexte.

Soit $f: M \to M'$ un homomorphisme de A- modules. On définit le *noyau* de f comme étant le noyau de cette application considérée comme homomorphisme de groupes additifs.

Par analogie avec des résultats précédents, on a:

Soit $f: M \to M'$ un homomorphisme de A-module. Le noyau et l'image de f sont alors des sous-modules de M et M' respectivement.

Démonstration. Soit E le noyau de f. On sait déjà que E est un sous-groupe additif de M. Soient $v \in M$ et $x \in A$. Alors

$$f(xv) = xf(v) = x0 = 0,$$

de sorte que $xv \in E$, et cela prouve que le noyau de f est un sous-module de M. On sait déjà que l'image de f est un sous-groupe de M'. Soit v' un élément de l'image de f, et $x \in A$. Soit v un élément de M tel que

$$f(v) = v'$$

Alors $f(xv) = xf(v) = xv'$ appartient aussi à l'image de M, qui est par conséquent un sous-module de M', prouvant ainsi notre assertion.

Exemple 5. Soit A un anneau, et M un idéal à gauche de A. Soit $y \in M$. L'application

$$r_y : M \to M$$

telle que

$$r_y(x) = xy,$$

est A-linéaire de M dans lui-même. En effet, si $x \in M$, alors $xy \in M$ puisque M est un idéal à gauche, et les conditions de A-linéaire sont ici des reformulations des définitions. Par exemple

$$\begin{aligned} r_y(x_1 + x_2) &= (x_1 + x_2)y = x_1y + x_2y \\ &= r_y(x_1) + r_y(x_2). \end{aligned}$$

De plus, pour $z \in A$ et $x \in M$

$$r_y(zx) = zxy = zr_y(x).$$

On appelle r_y la *multiplication à droite* par y. Ainsi r_y est un A-endomorphisme de M.

Remarquons que tout groupe abélien peut être considéré comme module sur les entiers. Ainsi un A-module M est aussi un $\mathbf{Z}$-module et tout A-endomorphisme de M aussi un endomorphisme de M considéré comme groupe abélien et $\mathrm{End}_A(M)$ un sous-ensemble de $\mathrm{End}(M) = \mathrm{End}_{\mathbf{Z}}(\mathrm{M})$.

En fait, $\mathrm{End}_A(M)$ est un sous-anneau de $\mathrm{End}(M)$, de sorte que $\mathrm{End}_A(M)$ est lui-même un anneau. La démonstration en est routinière. Par exemple, si $f, g \in \mathrm{End}_A(M)$ et si $x \in A$, $v \in M$, alors

$$\begin{aligned} (f + g)(xv) &= f(xv) + g(xv) \\ &= xf(v) + xg(v) \\ &= x\big(f(v) + g(v)\big) \\ &= x(f + g)(v). \end{aligned}$$

Donc, $f + g \in \mathrm{End}_A(M)$. Tout aussi facilement,

$$(f \circ g)(xv) = f\big(g(xv)\big) = f\big(xg(v)\big) = xf\big(g(v)\big).$$

L'identité est dans $\mathrm{End}_A(M)$. Cela prouve que $\mathrm{End}_A(M)$ est un sous-anneau de $\mathrm{End}_Z(M)$.

Nous voyons maintenant également que l'on peut considérer M comme module sur $\mathrm{End}_A(M)$, puisque M est un module sur $\mathrm{End}_Z(M) = \mathrm{End}(M)$.

Désignons $\mathrm{End}_A(M)$ par $A'(M)$, ou simplement par A' pour clarifier les notations. Soient $f \in A'$ et $x \in A$. On a, par définition,

$$f(xv) = xf(v),$$

et subséquemment,

$$f \circ \lambda_x(x) = \lambda_x \circ f(v).$$

L'application λ_x est, par suite, A'-linéaire de M dans lui-même, i.e. élément de $\mathrm{End}_{A'}(M)$.

La correspondance

$$\lambda : x \mapsto \lambda_x$$

est donc un homomorphisme d'anneaux de A dans $\mathrm{End}_{A'}(M)$, et pas seulement dans $\mathrm{End}(M)$.

***Théorème* 5.** *Soient A un anneau et M un A-module. Soit J l'ensemble des éléments $x \in A$ tels que $xv = 0$, pour tout $v \in M$. Alors J est un idéal bilatère de A.*

Démonstration. Si $x, y \in J$, alors $(x + y)v = xv + yv = 0$ pour tout $v \in M$. Si $a \in A$, alors

$$(ax)v = a(xv) = 0 \quad \text{et} \quad (xa)v = x(av) = 0,$$

pour tout $v \in M$. Cela démontre le théorème.

Remarquons que l'idéal bilatère du théorème 5 n'est autre que le noyau de l'homomorphisme d'anneaux

$$x \mapsto \lambda_x$$

décrit dans la discussion précédente.

***Théorème* 6.** (Wedderburn-Rieffel). *Soient A un anneau et L un idéal à gauche non nul considéré comme A-module. Soient $A' = \mathrm{End}_A(L)$ et $A'' = \mathrm{End}_{A'}(L)$. Soit*

$$\lambda : A \to A''$$

un homomorphisme d'anneau tel que $\lambda_x(y) = xy$, pour $x \in A$ et $y \in L$. Supposons que A n'ait pas d'autre idéal bilatère que 0 et A lui-même. Alors λ est un isomorphisme d'anneau.

Démonstration (Rieffel). Le fait que λ est injective vient du théorème 5 et de l'hypothèse que L n'est pas l'idéal nul. Par conséquent, la seule chose à démontrer est que λ est surjective. On sait d'après l'exemple 6 du chapitre III, §2, que LA

est un idéal bilatère, non nul puisque A possède un élément inversible, et est, par suite, égal à A, par hypothèse. Alors

$$\lambda(LA) = \lambda(L)\lambda(A) = \lambda(A).$$

Nous affirmons maintenant, que $\lambda(L)$ est un idéal à gauche de A''. Pour le démontrer, considérons $f \in A''$ et $x \in L$. Pour tout $y \in L$, nous savons, d'après l'exemple 5, que r_y est dans A', et que, par suite

$$f \circ r_y = r_y \circ f.$$

Cela signifie que $f(xy) = f(x)y$. On peut récrire cette relation sous la forme

$$f \circ \lambda_x(y) = \lambda_{f(x)}(y).$$

De là vient le fait que $f \circ \lambda_x$ est un élément de $\lambda(L)$, à savoir $\lambda_{f(x)}$. Cela prouve que $\lambda(L)$ est un idéal à gauche de A''. Mais alors

$$A''\lambda(A) = A''\lambda(L)\lambda(A) = \lambda(L)\lambda(A) = \lambda(A).$$

Puisque $\lambda(A)$ contient l'application identique, que nous appelons e, il s'ensuit que, pour tout $f \in A''$, l'application $f \circ e = f$ appartient à $\lambda(A)$, i.e. que A'' est contenu dans $\lambda(A)$, et donc que $A'' = \lambda(A)$, comme il fallait le démontrer.

L'importance du théorème 6 réside en ce qu'il permet de représenter A comme anneau d'endomorphismes d'un certain module, à savoir l'idéal à gauche L Cela est essentiel dans le cas suivant.

Soit D un corps non nécessairement commutatif i.e. un anneau dont l'ensemble des éléments non nuls est un groupe multiplicatif (et ainsi qu'en particulier, $1 \neq 0$ dans l'anneau). (Un tel anneau est parfois dit *corps gauche* - N.d.T.)

Soit A un anneau et M un module sur A. On dira que M est un *module simple* si $M \neq \{0\}$, et si M n'a pas d'autre sous-module que $\{0\}$ et M lui-même.

***Théorème* 7.** (Lemme de Schur.) *Soit M un module simple sur un anneau A. Alors* $\mathrm{End}_A(M)$ *est un corps (non nécessairement commutatif).*

Démonstration. Nous savons que $\mathrm{End}_A(M)$ est un anneau, et nous devons démontrer que tout élément non nul possède un inverse. Puisque $f \neq 0$, l'image de f est un sous-module de M qui n'est pas égal à M, de sorte que le noyau de f est $\{0\}$, et que f est, par conséquent, injective. Par suite, f possède un inverse comme homomorphisme de groupes, et on vérifie immédiatement que cet inverse est un A-homomorphisme, prouvant par là notre théorème.

Exemple 6. Soient A un anneau, et L un idéal à gauche qui est simple, en tant que A-module. (On dit que L est un idéal à gauche simple.) Alors $\mathrm{End}_A(L) = D$ est un corps. Si on est dans le cas où D est commutatif, alors, d'après le théorème 6, on conclut que $A = \mathrm{End}_D(L)$ est l'anneau de toutes les applications D-linéaires de L dans lui-même, et que L est un espace vectoriel sur le corps D. Ainsi nous avons une image concrète de l'anneau A.

Exercices

1. Soit A un anneau. Montrez que A peut être considéré comme module sur lui-même à un seul générateur.

2. Soient A un anneau et M un A-module. Montrez que $\mathrm{Hom}_A(A, M)$ et M sont isomorphes en tant que groupes additifs, par l'application $f \mapsto f(1)$.

3. Soient E et F deux A-modules. Montrez que $\mathrm{Hom}_A(E, F)$ est un module sur $\mathrm{End}_A(F)$, l'opération de l'anneau $\mathrm{End}_A(F)$ sur le groupe additif $\mathrm{Hom}_A(E, F)$ étant la composition des applications.

4. Soit E un module sur l'anneau A, et soit L un idéal à gauche de A. Soit LE l'ensemble de tous les éléments $x_1v_1 + \cdots + x_nv_n$ où $x_i \in A$ et $v_i \in E$. Montrez que LE est un sous-module de E.

5. (a) Soient A un anneau, E un module et L un idéal à gauche. Supposons que L et E soient simples. Montrez que $LE = E$, ou que $LE = 0$.

(b) Supposons que $LE = E$. Définissez la notion d'isomorphisme de modules. Démontrez que L est isomorphe à E, en tant que A-module. [*Indication:* soit $v_0 \in E$ un élément tel que $Lv_0 \neq \{0\}$. Montrez que l'application $x \mapsto xv_0$ établit un A-isomorphisme entre L et E.]

6. Soit A un anneau, E et F des A-modules. Soit $f: E \to F$ un isomorphisme. Montrez que $\mathrm{End}_A(E)$ et $\mathrm{End}_A(F)$ sont isomorphes en tant qu'anneaux, par l'application

$$f \mapsto \sigma \circ f \circ \sigma^{-1}$$

où $f \in \mathrm{End}_A(E)$.

7. Soient E et F des modules simples sur un anneau. Soit $f: E \to F$ un homomorphisme. Montrez que $f = 0$, ou que f est un isomorphisme.

8. Vérifiez en détail la dernière assertion faite dans la démonstration du théorème 7.

9. Soient A un anneau et E un module. On dit que E est un module *libre* s'il existe des éléments $v_1, \ldots, v_n$ dans E tels que tout élément $v \in E$ possède une décomposition unique

$$v = x_1v_1 + \cdots + x_nv_n$$

où les $x_i \in A$. Si c'est le cas, alors on dit que $v_1, \ldots, v_n$ est une base de E (sur A).

10. Soient E un module libre sur l'anneau A, de base $v_1, \ldots, v_n$, F un module et $w_1, \ldots, w_n$ des éléments de F. Montrez qu'il existe un unique homomorphisme $f: E \to F$ tel que $f(v_i) = w_i$ pour $i = 1, \ldots, n$.

11. Soient A un anneau, et E un ensemble constitué de n éléments, par exemple de $s_1, \ldots, s_n$. Soit F l'ensemble des applications de E dans A. (a) Montrez que F est un module. (b) Si $x \in A$, on désigne par xs_i la fonction de E dans A, qui associe x à s_i et 0 à s_j, pour $j \neq i$. Montrez que F est un module libre, que $1s_1, \ldots, 1s_n$ est une base de F sur A, et que tout élément s'exprime de façon unique sous la forme $x_1s_1 + \cdots + x_ns_n$, où $x_i \in A$.

12. Soient un anneau A, un module E et un sous-module F de E. Décrivez la façon de faire du groupe quotient E/F un A-module.

13. Soient K un corps et $A = K[X]$ l'anneau des polynômes à coefficients dans K. Soit J l'idéal engendré par X^2. Montrez que A/J est un K-espace vectoriel. Quelle est sa dimension?

14. Soient K un corps et $A = K[X]$ l'anneau des polynômes à coefficients dans K. Soit $f(X)$ un polynôme de degré $d > 0$ de $K[X]$. Soit J l'idéal engendré par $f(X)$. Quelle est sa dimension de A/J sur K? Exhibez une base du K-espace vectoriel A/J. Montrez que A/J est un anneau intègre si et seulement si f est irréductible.

15. Si A est un anneau commutatif, E et F des modules, montrez que $\mathrm{Hom}_A(E, F)$ est un A-module d'une façon naturelle. Est-ce encore vrai si A n'est pas commutatif?

16. Soient K un corps et A un espace vectoriel sur K de dimension 2. Soit $\{e, u\}$ une base de A. Si a, b, c, et d sont des éléments de K, on définit le produit

$$(ae + bu)(ce + du) = ace + (bc + ad)u.$$

Montrez que ce produit fait de A un anneau. Quel en est l'élément unité. Montrez que cet anneau est isomorphe à l'anneau $K[X]/(X^2)$ de l'exercice 13.

17. On conserve les notations de l'exercice précédent. Soit $f(X)$ un polynôme de $K[X]$. Montrez que

$$f(ae + u) = f(a)e + f'(a)u,$$

où f' est la dérivée formelle de f.

18. Soit A un anneau, et soient E', E et F des A-modules. Si $f : E' \to E$ est un A-homomorphisme, montrez que l'application $\varphi \mapsto f \circ \varphi$ est un $\mathbf{Z}$-homomorphisme de $\mathrm{Hom}_A(F, E)$ dans $\mathrm{Hom}_A(F, E)$, et est un A-homomorphisme si A est commutatif.

19. Une suite d'homomorphismes de groupes abéliens

$$A \xrightarrow{f} B \xrightarrow{g} C$$

est dite *exacte* si $\mathrm{Im}\, f = \mathrm{Ker}\, g$. Ainsi dire que $0 \longrightarrow A \xrightarrow{f} B$ est exacte signifie que f est injective. Soit A un anneau. Si

$$0 \longrightarrow E' \xrightarrow{f} E \xrightarrow{g} E''$$

est une suite exacte de A-modules, montrez que

$$0 \longrightarrow \mathrm{Hom}_A(F, E') \longrightarrow \mathrm{Hom}_A(F, E) \longrightarrow \mathrm{Hom}_A(F, E'')$$

est une suite exacte.

CHAPITRE VI
Théorie des corps

Nous supposons dans ce chapitre, et pour être concret, que tous les corps sont des sous-corps du corps des nombres complexes. En fait, les résultats restent valables si, au lieu de **C**, *on prend n'importe quel corps algébriquement clos et qui contient l'ensemble des rationnels comme sous-corps.*

§1. *Extensions algébriques*

Soit un corps K. On dit qu'un nombre α est *algébrique* sur K s'il existe un polynôme $f(t)$ non nul de $K[t]$ tel que $f(\alpha) = 0$, i.e. si α satisfait une équation polynomiale

$$a_n\alpha^n + \cdots + a_0 = 0$$

à coefficients dans K, non tous nuls. Si K est un sous-corps de E, et si tout élément de E est algébrique sur K, on dit que E est *algébrique* sur K.

Exemple 1. Si $\alpha^2 = 2$, i.e. si α est l'une des deux racines carrés possibles de 2, alors α est algébrique sur le corps **Q** des nombres rationnels. Une racine cubique de 2 est, de la même façon, algébrique. N'importe lequel des nombres $e^{2\pi i/n}$ (où n est un entier $\geqq 1$) est algébrique sur **Q**, puisque racine de $t^n - 1$. Il est bien connu (mais difficile à démontrer) que ni e, ni π ne sont algébriques sur **Q**.

Soit K un sous-corps d'un corps E. On peut considérer E comme espace vectoriel sur K. On dit aussi que E est une *extension* de K. On dira que E est une extension *finie* de K, si E est un espace vectoriel de dimension finie sur K. Par exemple, **C** est une extension finie de **R**, et $\{1, i\}$ une base de **C** sur **R**. Le corps des nombres réels ne constitue pas une extension finie de **Q**.

***Théorème* 1.** *Si E est une extension finie de K, alors tout élément de E est algébrique sur K.*

Démonstration. Les puissances $1, \alpha, \alpha^2, \ldots, \alpha^n$ d'un élément de E ne peuvent être algébriquement indépendants sur K, si $n > \dim E$. Il existe donc des éléments $a_0, \ldots, a_n \in K$, non tous nuls, tels que $a_n\alpha^n + \cdots + a_0 = 0$. Cela veut dire que α est algébrique sur K.

Soit α un nombre algébrique sur K. Soit J l'idéal des polynômes de $K[t]$ dont α est racine, i.e. *des polynômes f tels que $f(\alpha) = 0$. Soit $p(t)$ un générateur de J, de*

coefficient dominant égal à 1. *Alors p est irréductible. Démonstration:* Supposons que $p = gh$, avec $\deg g < \deg p$ et $\deg h < \deg p$. Puisque $p(\alpha) = 0$, on a $g(\alpha) = 0$ ou $h(\alpha) = 0$. Soit par exemple $g(\alpha) = 0$. Puisque $\deg g < \deg p$, cette situation est impossible à cause de l'hypothèse faite sur p.

Le polynôme irréductible p (de coefficient dominant 1) est uniquement déterminé par α dans $K[t]$, et on l'appelle *le polynôme irréductible* (ou *minimal*, N.d.T.) *de α sur K*. Son degré est appelé *degré* de α sur k. Nous allons immédiatement donner une autre interprétation de son degré.

***Théorème* 2.** *Soit α un nombre algébrique sur K. Soit n le degré de son polynôme irréductible sur K. L'espace vectoriel engendré sur K par 1, $\alpha, \ldots, \alpha^{n-1}$ est alors un corps, et la dimension de cet espace vectoriel est n.*

Démonstration. Soit f un polynôme quelconque de $K[t]$. On peut trouver q et r dans $K[t]$ tels que $f = qp + r$, avec $\deg r < \deg p$. Alors

$$f(\alpha) = q(\alpha)p(\alpha) + r(\alpha) = r(\alpha).$$

Par suite, en désignant par E l'espace vectoriel engendré par $1, \alpha, \ldots, \alpha^{n-1}$, on trouve que le produit de deux éléments de E est encore dans E. Supposons que $f(\alpha) \neq 0$. Le polynôme f n'est alors pas divisible par p. Il existe, par conséquent, des polynômes g, $h \in K[t]$ tels que

$$gf + hp = 1.$$

On obtient $g(\alpha)f(\alpha) + h(\alpha)p(\alpha) = 1$, d'où $g(\alpha)f(\alpha) = 1$. Ainsi tout élément non nul de E est inversible, et de là vient le fait que E soit un corps.

Le corps engendré par les puissances de α sur K, comme dans le théorème 2, sera désigné par $K(\alpha)$.

Si E est une extension finie de K, on désigne par

$$[E : K]$$

la dimension de E considérée comme espace vectoriel sur K, et on l'appelle le *degré* de E sur K.

***Théorème* 3.** *Soit E_1 une extension finie de K, et soit E_2 une extension finie de E_1. Alors E_2 est une extension finie de K, et*

$$[E_2 : K] = [E_2 : E_1][E_1 : K].$$

Démonstration. Soit $\{\alpha_1, \ldots, \alpha_n\}$ une base de E_1 sur K, et $\{\beta_1, \ldots, \beta_m\}$ une base de E_2 sur E_1. Démontrer que les éléments $\{\alpha_i\beta_j\}$ forment une base de E_2 sur K. Soit v un élément de E_2. On peut écrire

$$v = \sum_j w_j\beta_j = w_1\beta_1 + \cdots + w_m\beta_m$$

pour des $w_j \in E_1$. Ecrivons tous les w_j comme combinaisons linéaires de $\alpha_1, \ldots, \alpha_n$ à coefficients dans K, de sorte que

$$w_j = \sum_i c_{ij}\alpha_i.$$

On trouve, par substitution,

$$v = \sum_j \sum_i c_{ij}\alpha_i\beta_j.$$

Par suite les éléments $\alpha_i\beta_j$ engendrent E_2 sur K. Supposons que nous avons une relation

$$0 = \sum_j \sum_i x_{ij}\alpha_i\beta_j,$$

où les $x_{ij} \in K$. On a alors

$$\sum_j \left(\sum_i x_{ij}\alpha_i\right)\beta_j = 0.$$

De l'indépendance linéaire de $\beta_1, \ldots, \beta_m$ sur E_1, on tire

$$\sum_i x_{ij}\alpha_i = 0$$

pour tout j, et, de l'indépendance linéaire de $\alpha_1, \ldots, \alpha_n$ sur K, on tire $x_{ij} = 0$ pour tout i et tout j, comme il fallait le montrer.

Soient α et β deux nombres algébriques sur K. Le nombre β est, *a fortiori*, algébrique sur $K(\alpha)$. Nous pouvons former le corps $K(\alpha)(\beta)$. Tout corps contenant K, et α, β contiendra $K(\alpha)(\beta)$. Par suite, $K(\alpha)(\beta)$ est le plus petit corps contenant K et α et β tous les deux. De plus, d'après le théorème 3, $K(\alpha)(\beta)$ est finie sur K, se décomposant en la suite

$$K \subset K(\alpha) \subset K(\alpha)(\beta).$$

Par suite, d'après le théorème 3, le corps $K(\alpha)(\beta)$ est algébrique sur K. De plus, comme il est sans importance d'écrire $K(\alpha)(\beta)$ ou $K(\beta)(\alpha)$, nous désignons ce corps par $K(\alpha, \beta)$.

On constate, par récurrence, que, si $\alpha_1, \ldots, \alpha_r$ sont algébriques sur K et si $K(\alpha_1, \ldots, \alpha_r)$ est le plus petit corps contenant K et $\alpha_1, \ldots, \alpha_r$, on peut exprimer $K(\alpha_1, \ldots, \alpha_r)$ comme $K(\alpha_1)(\alpha_2) \cdots (\alpha_r)$. Ce dernier corps est algébrique sur K comme on peut le constater par application successive du théorème 3. Nous l'appelons le corps obtenu par *adjonction* à K des éléments $\alpha_1, \ldots, \alpha_r$.

Théorème 4. *Soit p un polynôme irréductible sur le corps K, de degré n. Le polynôme p possède alors n racines distinctes dans le corps des nombres complexes.*

Démonstration. On peut écrire

$$p(t) = (t - \alpha_1) \cdots (t - \alpha_n)$$

où les $\alpha_i \in \mathbf{C}$. Soit α une racine de p. Il va suffire de montrer que α est de multiplicité 1. Remarquons que p est le polynôme irréductible de α sur K. Remarquons aussi que la dérivée formelle p' de p est de degré $< n$. (cf. chapitre IV, §3.) Par suite,

nous ne pouvons pas avoir $p'(\alpha) = 0$, puisque p' n'est pas le polynôme nul (ce qui résulte immédiatement de la définition de la dérivée formelle, le coefficient dominant de p' étant $na_n \neq 0$). Finalement, α est de multiplicité 1.

Exercices

1. Soit $\alpha^2 = 2$. Montrez que le corps $\mathbf{Q}(\alpha)$ est de degré 2 sur $\mathbf{Q}$.
2. Montrez que le polynôme $(t - a)^2 + b^2$, où a et b sont rationnels, $b \neq 0$, est irréductible sur le corps des nombres rationnels.
3. Montrez que le polynôme $t^3 - p$ est irréductible sur $\mathbf{Q}$, pour tout entier p premier.
4. Quels sont les degrés des corps suivants sur $\mathbf{Q}$?
   (a) $Q(\alpha)$, où $\alpha^3 = 2$,
   (b) $Q(\alpha)$, où $\alpha^3 = p$ (p premier),
   (c) $\mathbf{Q}(\alpha)$, où α est racine de $t^3 - t - 1$,
   (*d*) $\mathbf{Q}(\alpha, \beta)$, où α est racine de $t^2 - 2$ et β de $t^2 - 3$?
5. Montrez que la racine cubique de l'unité $\omega = e^{2\pi i/3}$ est racine d'un polynôme de degré 2 sur $\mathbf{Q}$. Montrez que $\mathbf{Q}(\omega) = \mathbf{Q}(\sqrt{-3})$.
6. Quel est le degré du nombre $\cos(2\pi/3)$ sur $\mathbf{Q}$?
7. Quel est le degré du corps $\mathbf{Q}(i, \sqrt{3})$ sur $\mathbf{Q}$?
8. Soit E une extension de degré 2 d'un corps K. Montrez qu'on peut exprimer E sous la forme $K(\alpha)$, pour une certaine racine α d'un polynôme $t^2 - a$, où $a \in K$. [*Indication:* utilisez les formules de l'enseignement secondaire qui donnent les solutions des équations du second degré.]
9. Soit $t^2 + bt + c$ un polynôme de degré 2, où b et c sont dans K. Soit α une racine de ce polynôme. Montrez que $K(\alpha)$ est de degré 2 sur K si $b^2 - 4ac$ n'est pas un carré dans K, et que, sinon, $K(\alpha)$ est de degré 1 sur K, i.e. que $\alpha \in K$.
10. Soit $a \in \mathbf{C}$, et $a \neq 0$. Soit α une racine de $t^n - a$. Montrez que toutes les racines de $t^n - a$ sont du type $\omega\alpha$, où ω est une racine n-ième de l'unité, i.e.

$$\omega = e^{2\pi ik/n}, \qquad \text{pour} \quad k = 0, \ldots, n - 1.$$

§2. *Plongements*

Soit K un corps, et soit L un autre corps. Par *plongement* de K dans L, on entend une application

$$\sigma : K \to L$$

telle que, pour tous $x, y \in K$, on a

$$\sigma(x + y) = \sigma(x) + \sigma(y) \quad \text{et} \quad \sigma(xy) = \sigma(x)\sigma(y),$$

et telle que $\sigma(1) = 1$. Il s'ensuit alors que, si $x \neq 0$, $\sigma(x) \neq 0$ (car $1 = \sigma(xx^{-1}) = \sigma(x)\sigma(x^{-1})$). Par conséquent σ est un homomorphisme à la fois pour le groupe additif de K et pour le groupe multiplicatif des éléments non nuls de K. De plus, puisque le noyau de σ considéré comme homomorphisme additif, est $\{0\}$, il s'ensuit que σ est injectif, i.e. que $\sigma(x) \neq \sigma(y)$ si $x \neq y$. C'est la raison pour laquelle σ est appelé plongement. Nous écrivons souvent σx au lieu de $\sigma(x)$, et σK au lieu de $\sigma(K)$.

Un plongement $\sigma : K \to K'$ est dit *isomorphisme* si l'image de σ est K'. (On devrait spécifier qu'il s'agit d'un *isomorphisme de corps*, mais le contexte rend toujours clair ce que nous voulons dire.) Si $\sigma : K \to L$ est un plongement, alors l'image σK de K par σ est évidemment un sous-corps de L, et σ fournit ainsi un isomorphisme de K avec σK. Si $\sigma : K \to K'$ est un isomorphisme, on peut définir un isomorphisme inverse $\sigma^{-1} : K' \to K$ de la façon habituelle.

Soit $f(t)$ un polynôme de $K[t]$. Soit σ un plongement de K dans L. Posons

$$f(t) = a_n t^n + \cdots + a_0,$$

et définissons σf comme étant le polynôme

$$\sigma f(t) = \sigma(a_n) t^n + \cdots + \sigma(a_0).$$

On vérifie alors trivialement que, pour deux polynômes f et g de $K[t]$, on a

$$\sigma(f + g) = \sigma f + \sigma g \quad \text{et} \quad \sigma(fg) = (\sigma f)(\sigma g).$$

Si $p(t)$ est un polynôme irréductible de $K[t]$, alors p est irréductible sur K. C'est là un fait important. Il est facile de le démontrer, car, si nous avons une décomposition

$$p = gh$$

sur σk, alors

$$p = \sigma^{-1}\sigma p = (\sigma^{-1}g)(\sigma^{-1}h)$$

se factorise sur K.

Soit $f(t) \in k(t)$, et soit α un nombre algébrique sur K. Soit un plongement σ de $K(\alpha)$ dans un corps K. On a alors

$$(\sigma f)(\sigma\alpha) = \sigma\big(f(\alpha)\big).$$

Cela résulte immédiatement de la définition d'un plongement, car si $f(t)$ est le même que ci-dessus,

$$f(\alpha) = a_n\alpha^n + \cdots + a_0,$$

d'où

(*) $$\sigma\big(f(\alpha)\big) = \sigma(a_n)\sigma(\alpha)^n + \cdots + \sigma(a_0).$$

En particulier, si α est racine de f, $\sigma(\alpha)$ est racine de σf. Remarquons aussi que si σ est un plongement de $K(\alpha)$ dont l'effet est connu sur K et sur α, alors l'effet de σ est déterminé de façon unique sur $K(\alpha)$ par (*).

Soit $\sigma : K \to L$ un plongement. Soit E une extension de K. Un plongement $\tau : E \to L$ est dit *prolongement* de σ si $\tau(x) = \sigma(x)$ pour tout $x \in K$. On dit aussi que σ est une *restriction* de τ à K.

Théorème 5. *Soit $\sigma : K \to L$ un plongement. Soit $p(t)$ un polynôme irréductible $K[t]$. Soient α une racine de p, et β une racine de σp dans L. Il existe alors*

un plongement $\tau : K(\alpha) \to L$ qui prolonge σ, et tel que $\tau\alpha = \beta$. Réciproquement, tout prolongement τ de σ à $K(\beta)$ est tel que $\tau\alpha$ est racine de σp.

Démonstration. La seconde assertion résulte d'une remarque que nous avons faite précédemment. Pour prouver l'existence de τ, considérons un polynôme f quelconque de $K[t]$ et définissons τ par l'image $(\sigma f)(\beta)$ qu'il donne de l'élément $f(\alpha)$. Le même élément $f(\alpha)$ possède un grand nombre de représentations sous la forme $f(\alpha)$, pour un grand nombre de polynômes f de $K[t]$. Nous devons donc montrer que notre définition de τ ne dépend pas du choix de f. Supposons que f et g, appartenant à $K[t]$, soient tels que $f(\alpha) = g(\alpha)$. On a alors $(f - g)(\alpha) = 0$. Par suite, il existe un polynôme h de $K[t]$ tel que $f - g = ph$. Alors

$$\sigma f = \sigma g + (\sigma p)(\sigma h).$$

Par suite

$$\begin{aligned}(\sigma f)(\beta) &= (\sigma g)(\beta) + (\sigma p)(\beta) \cdot (\sigma h)(\beta) \\ &= (\sigma g)(\beta).\end{aligned}$$

Cela prouve que nous avons bien défini une application. Nous avons essentiellement utilisé le fait que p est irréductible! Il est maintenant trivial de vérifier que τ est un plongement, et nous laissons cela au lecteur.

Corollaire 1. *Soit p un polynôme irréductible sur le corps K. Soit α une racine de p. Soit*

$$\sigma : K \to \mathbf{C}$$

un plongement de K dans le corps des nombres complexes. Le nombre de plongements possibles de $K(\alpha)$ dans $\mathbf{C}$ qui prolongent σ est alors égal au degré de p (i.e. *au degré de α sur K*).

Démonstration. Conséquence immédiate des théorèmes 4 et 5.

Corollaire 2. *Soit E une extension finie de K. Soit n le degré de E sur K. Soit $\sigma : K \to \mathbf{C}$ un plongement de K dans le corps des nombres complexes. Le nombre de prolongements de σ à un plongement de E dans $\mathbf{C}$ est égal à n.*

Preuve. On peut écrire E sous la forme $E = K(\alpha_1, ..., \alpha_r)$. Considérons la suite

$$K \subset K(\alpha_1) \subset K(\alpha_1, \alpha_2) \subset \cdots \subset K(\alpha_1, ..., \alpha_r)$$

Soit $E_{r-1} = K(\alpha_1, ..., \alpha_{r-1})$. Supposons que nous ayons démontré par récurrence que le nombre de prolongements de σ à E_{r-1} est égal au degré $[E_{r-1} : K]$. Soient $\sigma_1, ..., \sigma_m$ les prolongements de σ à E_{r-1}. Soit d le degré de α_r sur E_{r-1}. Pour chaque $i = 1, ..., m$, on peut trouver exactement d prolongements de σ_i à E; appelons-les $\sigma_{i1}, ..., \sigma_{id}$. Il est clair alors que l'ensemble $\{\sigma_{ij}\}$ ($i = 1, ..., m$ et $j = 1, ..., d$) est l'ensemble des prolongements distincts de σ à E. Cela prouve notre corollaire.

Soit α un élément algébrique sur K. Soit $p(t)$ le polynôme irréductible de α sur K. Soient $\alpha_1, ..., \alpha_n$ les racines de p. Nous appelons ces racines les *conjuguées* de α sur K. Pour chaque α_i, il existe un plongement σ_i de K(α) qui envoie α sur α_i et qui est l'identité sur K. Ce plongement est uniquement déterminé.

Exemple 1. Considérons une racine α du polynôme $t^3 - 2$. Prenons pour α la racine cubique réelle de 2, écrite $\alpha = \sqrt[3]{2}$. Soient $1, \omega\ \omega^2$ les trois racines cubiques de l'unité. Le polynôme $t^3 - 2$ est irréductible sur $\mathbf{Q}$, parce qu'il n'a pas de racine dans $\mathbf{Q}$ (cf. exercices 1, 2 du chapitre IV, §3). Il existe par suite trois plongements de $\mathbf{Q}(\alpha)$ dans $\mathbf{C}$, à savoir les trois plongements $\sigma_1, \sigma_2, \sigma_3$ tels que

$$\sigma_1\alpha = \alpha, \quad \sigma_2\alpha = \omega\alpha, \quad \sigma_3\alpha = \omega^2\alpha$$

Exemple 2. Si $\alpha = 1 + \sqrt{2}$, il existe deux plongements de $\mathbf{Q}(\alpha)$ dans $\mathbf{C}$, à savoir ceux envoyant respectivement α sur $1 + \sqrt{2}$ et $1 - \sqrt{2}$.

***Théorème* 6.** (*de l'élément primitif* - N.d.T.) *Soit E une extension finie de K. Il existe alors un élément γ de E tel que $E = K(\gamma)$.*

Démonstration. Il va suffire de démontrer que, si $E = K(\alpha, \beta)$ pour deux éléments α et β algébriques sur K, alors on peut trouver γ dans E tel que $E = K(\gamma)$, car alors on peut raisonner par récurrence. Soit $[E : F] = n$. Soient $\sigma_1, \ldots, \sigma_n$ les n plongements distincts de E dans $\mathbf{C}$, prolongeant l'identité sur K. Nous allons d'abord démontrer qu'on peut trouver un élément $c \in K$ tel que les éléments

$$\sigma_i\alpha + c\sigma_i\beta = \sigma_i(\alpha + c\beta)$$

sont distincts, pour $i = 1, \ldots, n$. Considérons les polynômes

$$\prod_{i=1}^{n} \prod_{j \neq i} [\sigma_j\alpha - \sigma_i\alpha + t(\sigma_j\beta - \sigma_i\beta)].$$

Ce n'est pas le polynôme nul, puisque tous les facteurs sont différents de 0. Ce polynôme possède un nombre fini de racines. Par suite, on peut certainement trouver un élément c de K qui, substitué à t, n'annule pas ce polynôme. Cet élément c va réaliser ce que nous cherchons à faire.

Nous affirmons en effet maintenant que $E = K(\gamma)$, où $\gamma = \alpha + c\beta$. En fait, par construction, nous avons n plongements distincts de $K(\gamma)$ dans $\mathbf{C}$ prolongeant l'identité sur K, à savoir $\sigma_1, \ldots, \sigma_n$. Par suite $[K(\gamma) : K] \geqq n$, d'après le corollaire 1 du théorème 5. Puisque $K(\gamma)$ est un sous-espace de E sur K qui a la même dimension que E, il en résulte que $K(\gamma) = E$, et notre théorème est démontré.

Exemple 3. Démontrez en exercice que si $\alpha^3 = 2$, et si β est une racine carrée de 2, alors $\mathbf{Q}(\alpha, \beta) = \mathbf{Q}(\gamma)$, où $\gamma = \alpha + \beta$.

Exercices

1. Trouvez dans tous les cas qui suivent un élément α tel que $\mathbf{Q}(\alpha, \beta) = \mathbf{Q}(\gamma)$. Démontrez toujours les affirmations que vous faites.

(a) $\alpha = \sqrt{-5}, \beta = \sqrt{2}$ (b) $\alpha = \sqrt[3]{2}, \beta = \sqrt{2},$
(c) $\alpha =$ racine de $t^3 - t + 1$, $\beta =$ racine de $t^2 - t - 1$,
(d) $\alpha =$ racine de $t^3 - 2t + 3$, $\beta =$ racine de $t^2 + t + 2$.

2. Déterminez les degrés des corps $\mathbf{Q}(\alpha, \beta)$ sur $\mathbf{Q}$ dans chacun des cas de l'exercice 2.

3. Soient E_1 et E_2 deux extensions d'un corps K. Supposons que $[E_2 : K] = 2$ et que $E_1 \cap E_2 = K$. Soit $E_2 = K(\alpha)$. Montrez que $E_1(\alpha)$ est de degré 2 sur E_1.

4. Soit $\alpha^3 = 2$, soit ω une racine cubique complexe de l'unité et soit $\beta = \omega\alpha$. Quel est le degré de $\mathbf{Q}(\alpha, \beta)$ sur $\mathbf{Q}$.

5. Soient E_1 de degré p et E_2 de degré p' sur K, pour des entiers p et p' premiers entre eux. Montrez que, soit $E_1 = E_2$, soit $E_1 \cap E_2 = K$.

6. Soit E une extension finie de K, de degré n. Soient $\sigma_1, \ldots, \sigma_n$ les plongements distincts de E sur K dans $\mathbf{C}$. On définit, pour un $\alpha \in E$, la *trace* et la *norme* de α (de E sur K) par, respectivement,

$$Tr_K^E(\alpha) = \sum_{i=1}^{n} \sigma_i\alpha = \sigma_1\alpha + \cdots + \sigma_n\alpha,$$

$$N_K^E(\alpha) = \prod_{i=1}^{n} \sigma_i\alpha = \sigma_1\alpha \ldots \sigma_n\alpha.$$

(a) Montrez que la trace et la norme de α sont dans K.

(b) Montrez que la trace est un homomorphisme additif, et la norme un homomorphisme multiplicatif.

7. Soit α un élément algébrique sur le corps K, et soit

$$p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0$$

le polynôme irréductible de α sur K. Montrez que

$$N(\alpha) = (-1)^n a_0 \qquad \text{et} \qquad Tr(\alpha) = -a_{n-1}.$$

(Les normes et les traces sont prises de $K(\alpha)$ sur K.)

8. Soit E une extension finie de K, et soit a un élément de K. Soit $[E : K] = n$. Quelles sont la norme et la trace de a de E sur K ?

§3. *Corps de dislocation*

Soit E une extension finie de K. Soit σ un plongement de K, et τ un prolongement de σ à un plongement de E. Nous disons également que τ est *sur* σ. Si σ est l'identité, nous disons alors que τ est un *plongement de E sur K* (ou un *K-plongement*-N.d.T.) Le plongement τ de E sur K est alors tel que $\tau x = x$ pour tout $x \in K$. On dit aussi que τ laisse *K fixe*.

Nous entendons par *automorphisme* d'un corps K un isomorphisme $\sigma : K \to K$ de K avec lui-même. Le contexte montre toujours clairement qu'il s'agit d'un isomorphisme de corps (et d'aucun autre type d'isomorphisme, comme ceux de groupes ou d'espaces vectoriels).

Soit σ un plongement sur K d'une extension finie L de K. Supposons que $\sigma(L)$ soit contenue dans L. Alors $\sigma(L) = L$. En effet, σ induit une application linéaire de l'espace vectoriel de L sur K, qui est injective. Par suite, σ est surjective et est donc un isomorphisme (de corps ou d'espace vectoriel), et par conséquent un automorphisme.

Remarquons que l'ensemble des automorphismes d'un corps L est un groupe. On le vérifie trivialement. Nous allons nous occuper de certains de ses sous-groupes.

Soit G un groupe d'automorphismes d'un corps L. Soit L^G l'ensemble de tous les éléments $x \in L$ tels que $\sigma x = x$ pour tout $\sigma \in G$. L'ensemble L^G est alors un corps.

En effet, L^G contient 0 et 1. Si x et y sont dans L^G, alors

$$\sigma(x + y) = \sigma x + \sigma y = x + y,$$
$$\sigma(xy) = \sigma(x)\sigma(y) = xy,$$

de sorte que $x + y$ et xy sont dans L^G. On a aussi $\sigma(x^{-1}) = \sigma(x)^{-1} = x^{-1}$, de sorte que x^{-1} est dans L^G. Cela prouve que L^G est un corps appelé *corps fixe* de G, (ou *corps des invariants* de G - N.d.T.).

Si G est un groupe de K-automorphismes de L sur un sous-corps K, alors K est contenu dans le corps fixe de G par définition, mais le corps fixe peut être plus grand que K. Si, par exemple, G est réduit à l'identité, alors son corps fixe est L lui-même.

Exemple 1. Le corps des nombres rationnels n'a pas d'autre automorphisme que l'identité. Démonstration?

Exemple 2. Démontrez que le corps $\mathbf{Q}(\alpha)$, où $\alpha^3 = 2$ n'a pas d'autre automorphisme que l'identité.

Exemple 3. Soit K un corps et $a \in K$. Supposons que a n'est pas un carré dans K; soit $\alpha^2 = a$. Alors $F(\alpha)$ possède exactement deux K-automorphismes, à savoir l'identité, et l'automorphisme qui envoie α sur $-\alpha$.

Une extension finie L de K sera dite *galoisienne* si tout K-plongement de L est un automorphisme de L.

Une extension finie L de K est dite *corps de dislocation* (ou *corps de rupture*. N.d.T.) d'un polynôme si $L = K(\alpha_1, ..., \alpha_n)$, où $\alpha_1, ..., \alpha_n$ sont les racines du polynôme.

***Théorème* 7.** *Une extension finie de K est galoisienne si et seulement si elle est corps de dislocation d'un polynôme.*

Démonstration. Soit L une extension galoisienne de K. Posons, en utilisant le théorème 6, $L = K(\alpha)$ pour un certain élément α. Soit $p(t)$ le polynôme irréductible de α sur K. Il existe, pour toute racine α_i de p, un unique K-plongement σ_i de L tel que $\sigma_i\alpha = \alpha_i$. Puisque tout plongement est un automorphisme, il s'ensuit que α_i est contenu dans L. Par suite

$$L = K(\alpha) = K(\alpha_1, ..., \alpha_n)$$

et L est le corps de dislocation de p.

Réciproquement supposons que K soit le corps de dislocation d'un polynôme $f(t)$, non nécessairement irréductible, dont les racines sont $\alpha_1, ..., \alpha_n$. Si σ est un K-plongement de L, alors $\sigma\alpha_i$ doit aussi être racine de K. Par suite σ envoie L dans lui-même, et est un automorphisme.

***Théorème* 8.** *Soit L une extension galoisienne de K. Si $p(t)$ est un polynôme de $K[t]$, s'il est irréductible sur K et si p possède une racine dans L, alors p a toutes ses racines dans L.*

Démonstration. Soit α l'une des racines de p dans L. Soit β une autre racine de p. D'après le théorème 5, il existe un $K(\alpha)$-plongement σ de $K(\beta)$ envoyant α sur β, et égal à l'identité sur K. Prolongeons ce plongement à L. Puisque un K-plongement de L est un automorphisme, on doit avoir $\sigma\alpha \in L$, et donc $\beta \in L$.

§4. *Théorème fondamental*

Théorème 9. *Soit L une extension galoisienne de K. Soit G le groupe des K-automorphismes de L. Alors K est le corps des invariants de G.*

Démonstration. Soit K' le corps fixe de G. On a trivialement $K \subset K'$. Supposons que $\alpha \in K'$ et $\alpha \notin K$. D'après le théorème 5, il existe un *K-plongement* σ_0 de $K(\alpha)$ tel que $\sigma_0\alpha \neq \alpha$. Prolongeons σ_0 à un K-plongement de L. (Corollaire 2 du théorème 5.) Par hypothèse, σ est un K-automorphisme de L, et $\sigma\alpha = \sigma_0\alpha \neq \alpha$, contredisant ainsi l'hypothèse que $\alpha \in K'$ mais que $\alpha \notin K$. Cela prouve notre théorème.

***Théorème* 10.** *Soit L une extension galoisienne de K. A tout corps intermédiaire E, on associe le sous-groupe $G_{L/E}$ des automorphismes de L laissant E fixe. L'extension L est galoisienne sur E, et l'application*

$$E \mapsto G_{L/E}$$

est une application injective et surjective de l'ensemble des corps intermédiaires, sur l'ensemble des sous-groupes de G, et E est le corps des invariants de $G_{L/E}$.

Démonstration. Tout E-plongement de L est un K-plongement, et par suite un automorphisme de K. Il en résulte que L est galoisienne sur E. De plus, E est le corps fixe de $G_{L/E}$ d'après le théorème 9. Cela montre, en particulier, que l'application

$$E \mapsto G_{L/E}$$

est injective, i.e. que si $E \neq E'$ alors $G_{L/E} \neq G_{L/E'}$. Soit enfin H un sous-groupe de G. On peut écrire $L = K(\alpha)$ pour un certain élément α. Soit $\{\sigma_1, ..., \sigma_r\}$ les éléments de H, et soit

$$f(t) = (t - \sigma_1\alpha) \cdots (t - \sigma_r\alpha).$$

Remarquons que $\{\sigma\sigma_1, ..., \sigma\sigma_r\}$ est une permutation de $\{\sigma_1, ..., \sigma_r\}$, pour tout $\sigma \in H$. Par suite, de l'expression

$$\sigma f(t) = (t - \sigma\sigma_1\alpha) \cdots (t - \sigma\sigma_r\alpha) = f(t),$$

on tire que les coefficients de f sont dans le corps fixe E de H. De plus $L = K(\alpha)$, et α est racine d'un polynôme de degré r sur E. Par suite, $[L : E] \leqq r$. Mais L possède r E-plongements distincts (ceux de H), et par conséquent, d'après le raisonnement habituel, $[L : E] = r$, et $H = G_{L/E}$. Cela prouve notre théorème.

Si L est une extension galoisienne de K, le groupe des automorphismes de $G_{L/K}$ est appelé *groupe de Galois* de L sur K. Si L est le corps de dislocation d'un

polynôme $f(t)$ dans $K[t]$, on dit aussi que $G_{L/K}$ est le *groupe de Galois de* f.

Soit $f(t) \in F(t)$, et soit

$$f(t) = (t - \alpha_1) \cdots (t - \alpha_n)$$

Soit $L = K(\alpha_1, \ldots, \alpha_n)$, et soit σ un élément de $G_{L/K}$. Alors $\{\sigma\alpha_1, \ldots, \sigma\alpha_n\}$ est une permutation de $\{\alpha_1, \ldots, \alpha_n\}$, que l'on peut noter π_σ. Si $\sigma \neq \tau$, $\pi_\sigma \neq \pi_\tau$, et, de façon claire,

$$\pi_{\sigma\tau} = \pi_\sigma \circ \pi_\tau .$$

Nous avons par conséquent représenté le groupe de Galois $G_{L/K}$ comme groupe de permutations des racines de f. Naturellement il n'est pas toujours vrai que l'on peut représenter toute permutation de $\{\alpha_1, \ldots, \alpha_n\}$ par un élément de $G_{L/K}$, même si f est irréductible sur K. Voyez les paragraphes suivants pour les exemples.

Exercices

1. On entend par racine *primitive* n-ième de l'unité, un nombre ω dont la période est exactement n. Par exemple, $e^{2\pi i/n}$ est une racine primitive n-ième de l'unité. Montrez que toute autre racine primitive n-ième de l'unité est égale à une puissance $e^{2\pi i r/n}$, où r est un entier > 0 et premier avec n.

2. Soient K un corps et $L = K(\omega)$, où ω est une racine primitive n-ième de l'unité. Montrez que L est galoisienne sur K, et que son groupe de Galois est commutatif. [*Indication*: remarquez que pour tout K-plongement σ, $\sigma\omega = \omega^{r(\sigma)}$ pour un certain entier $r(\omega)$.] Si τ est un autre plongement, qu'est-ce que $\tau\sigma\omega$, et $\sigma\tau\omega$.

3. Soient K_1 et K_2 deux extensions galoisiennes d'un corps K. Soient, pour fixer les idées, $K_1 = K(\alpha_1)$ et $K_2 = K(\alpha_2)$. Soit $L = K(\alpha_1, \alpha_2)$. Montrez que L est galoisienne sur K. Soit G son groupe de Galois. Montrez qu'on peut envoyer G dans le produit direct $G_{K_1/K} \times G_{K_2/K}$ en associant à tout σ de G le couple (σ_1, σ_2), où σ_1 est la restriction de σ à K_1, et σ_2 celle de σ à K_2. Montrez que cette application est un homomorphisme injectif.

4. Soit L une extension galoisienne de K, et soit E un corps intermédiaire ($K \subset E \subset L$), tel que E soit galoisienne sur K. Soit G le groupe de Galois de L sur K, et H le groupe de Galois de L sur E. Soit $\mathrm{res}_E\, \sigma$ la restriction à E de $\sigma \in G$. Montrez que l'application $\sigma \mapsto \mathrm{res}_E\, \sigma$ est un homomorphisme surjectif de G sur $G_{E/K}$ dont le noyau est H. Par suite, $G_{E/K}$ est isomorphe au groupe quotient G/H.

5. Soit K un corps contenant $i = \sqrt{-1}$. Soit L le corps de dislocation du polynôme $t^4 - a$, où $a \in K$. Montrez que le groupe de Galois de L sur K est un sous-groupe d'un groupe cyclique d'ordre 4. Si $t^4 - a$ est irréductible sur K, montrez que son groupe de Galois est cyclique d'ordre 4. Si α est une racine de $t^4 - a$, exprimez toutes les autres racines en fonction de α et de i.

6. Soit, plus généralement, un corps K contenant toutes les racines n-ièmes de l'unité. Soit L un corps de dislocation de l'équation $t^n - a = 0$, avec $a \in K$. Montrez que L est galoisienne sur K, et que son groupe de Galois est un sous-groupe d'un groupe cyclique d'ordre n.

7. Montrez que le groupe de Galois du polynôme $t^4 - 2$ sur le corps des rationnels est d'ordre 8, et contient un sous-groupe cyclique d'ordre 4. [*Indication*: démontrer d'abord que le polynôme est irréductible sur $\mathbf{Q}$. Considérer alors, si α est une racine quatrième réelle de 2, $L = \mathbf{Q}(\alpha, i)$.]

§5. *Extensions quadratiques et cubiques*

Résumons d'abord les propriétés des extensions quadratiques. Soit K un corps. Tout polynôme irréductible $t^2 + bt + c$ à coefficients dans K a $K(\alpha)$ pour corps de dislocation, où

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4c}}{2}.$$

L'extension $K(\alpha)$ est galoisienne sur K, et son groupe de Galois est cyclique d'ordre 2. Si on pose $d = b^2 - 4c$ alors $K(\alpha) = K(\sqrt{d})$. Réciproquement, le polynôme $t^2 - d$ est irréductible sur K, si et seulement si d n'est pas un carré dans K.

Considérons maintenant le cas cubique. Après élimination du terme carré, un polynôme cubique de $K[t]$ peut être mis sous la forme

$$f(t) = t^3 + bt + c = (t - \alpha_1)(t - \alpha_2)(t - \alpha_3),$$

où $b, c \in K$. Les racines peuvent, ou non, être dans K. Si f n'a pas de racine dans K, alors f est irréductible. On trouve

$$\alpha_1 + \alpha_2 + \alpha_3 = 0, \quad \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 = b, \quad -\alpha_1\alpha_2\alpha_3 = c.$$

Définissons le *discriminant* de f comme étant

$$D = [(\alpha_2 - \alpha_1)(\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2)]^2.$$

Tout automorphisme de $K(\alpha_1, \alpha_2, \alpha_3)$ laisse D invariant puisqu'il ne transforme le produit

$$\delta = (\alpha_2 - \alpha_1)(\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2)$$

qu'au plus en le changeant de signe.

Soit $L = K(\alpha_1, \alpha_2, \alpha_3)$ le corps de dislocation de f. Soit G le groupe de Galois de L sur K. Supposons que f est irréductible. On peut alors représenter G comme sous-groupe du groupe symétrique S_3. Puisque L contient $K(\alpha)$ pour toute racine α de f, il s'ensuit que $[L : K]$ est divisible par 3, et que par suite l'ordre de G est 3 ou 6. Dans le premier cas, $L = K(\alpha)$, et G est cyclique d'ordre 3. Dans le second cas, G est isomorphe à S_3. Démontrez en exercice le:

***Théorème* 11.** *Le groupe G est isomorphe à S_3 si et seulement si D n'est pas un carré dans K. Et si D est un carré dans K, l'extension L est de degré* 3 *sur K.*

Démontrez, en utilisant ce qui a déjà été dit, l'autre exercice facile suivant:

***Théorème* 12.** *Soit $f(t) = t^3 + at + b$ un polynôme irréductible sur K. le corps de dislocation de f. L'extension L est alors égale à $K(\sqrt{D}, \alpha)$ pour toute racine α de f.*

Puisque D est dans K, nous avons le sentiment qu'il doit y avoir moyen d'exprimer D en fonction de a et b. Vous trouverez, après quelques calculs, que

$$D = -4a^3 - 27b^2.$$

On peut alors, à l'aide de ce renseignement, déterminer explicitement, le groupe de Galois du polynôme cubique.

Soulignons qu'avant d'entreprendre quoi que ce soit d'autre, on doit *toujours* déterminer si f est irréductible ou non. Pour les polynômes cubiques, on peut le faire à l'aide des théorèmes 12 et 13 du chapitre IV, §5.

Exemple. Considérons le polynôme $f(t) = t^3 - 3t + 1$. Il n'a pas de racine entière, et est donc irréductible sur $\mathbf{Q}$. Son discriminant est

$$D = -4a^3 - 27b^2 = 3^4$$

Le discriminant est un carré dans $\mathbf{Q}$, et, par suite, le groupe de Galois de f sur $\mathbf{Q}$ est cyclique, d'ordre 3. Le corps de dislocation de f est $\mathbf{Q}(\alpha)$ pour n'importe laquelle des racines α de f.

Exercices

1. Déterminez les groupes de Galois des polynômes suivants sur le corps des nombres rationnels.

(a) $t^2 - t + 1$; (b) $t^2 - 4$; (c) $t^2 + t + 1$; (d) $t^2 - 27$.

2. Déterminez les groupes de Galois des polynômes suivants sur le corps des nombres rationnels. Trouvez-en les discriminants.

(a) $t^3 - 3t + 1$; (b) $t^3 + 3$; (c) $t^3 - 5$;
(d) $t^3 - a$, où a est rationnel, $\neq 0$, et n'est pas le cube d'un nombre rationnel;
(e) $t^3 - 5t + 7$; (f) $t^3 + 2t + 2$; (g) $t^3 - t - 1$.

3. Déterminez les groupes de Galois des polynômes suivants sur les corps indiqués:

(a) $t^2 - 10$ sur $\mathbf{Q}(\sqrt{2})$; (b) $t^3 - 10$ sur $\mathbf{Q}$;
(c) $t^3 - t - 1$ sur $\mathbf{Q}(\sqrt{-23})$; (d) $t^3 - 10$ sur $\mathbf{Q}(\sqrt{-3})$;
(e) $t^3 - 2$ sur $\mathbf{Q}(\sqrt{-3})$; (f) $t^3 - 9$ sur $\mathbf{Q}(\sqrt{-3})$;
(g) $t^2 - 5$ sur $\mathbf{Q}(\sqrt{-5})$; (h) $t^2 + 5$ sur $\mathbf{Q}(\sqrt{-5})$.

4. Soit $f(t) = t^3 + at + b$. Soit α une racine de f, et soit β un nombre tel que

$$\alpha = \beta - \frac{a}{3\beta}.$$

Montrez qu'on ne peut trouver un tel β si $a \neq 0$. (Pourquoi?) Montrez que

$$\beta^3 = +b/2 + \sqrt{+D/108}.$$

On obtient, de cette manière, une expression de α en termes de radicaux.

§6. *Résolubilité par radicaux*

Considérons deux cas particuliers caractéristiques du cas général, avant d'aborder le théorème fondamental.

Une extension galoisienne dont le groupe de Galois est abélien est dite *abélienne*.

Soit L une extension galoisienne de K, $L = K(\alpha)$. Soient τ et σ des K-automorphismes de L. Pour vérifier que $\sigma\tau = \tau\sigma$, il suffit de vérifier que $\sigma\tau\alpha = \tau\sigma\alpha$. En effet, on peut écrire tout élément de L sous la forme

$$x = a_0 + a_1\alpha + \cdots + a_{d-1}\alpha^{d-1},$$

si d est le degré de α sur K. Puisque $\sigma\tau a_i = \tau\sigma a_i$ pour tout i, il s'ensuit que, en plus de $\sigma\tau\alpha = \tau\sigma\alpha$, $\sigma\tau\alpha^i = \tau\sigma\alpha^i$ pour tout i, d'où $\tau\sigma x = \sigma\tau x$. Nous allons développer deux cas particuliers importants.

(1) *Soit K un corps et n un entier positif. Soit ω une racine n-ième primitive de l'unité, c'est-à-dire telle que $\omega^n = 1$ et que toute racine n-ième de l'unité puisse s'exprimer sous la forme ω^r pour un r vérifiant $0 \leqq r < n$. Soit $L = K(\omega)$. Nous allons démontrer que L est galoisienne et abélienne sur K.* Soit σ un K-plongement de L. On a alors

$$(\sigma\omega)^n = \sigma(\omega)^n = 1.$$

Par suite $\sigma\omega$ est aussi une racine n-ième de l'unité, et il existe un entier r tel que $\sigma\omega = \omega^r$. En particulier, L est galoisienne sur K. De plus, si τ est un autre automorphisme de L sur K, alors $\tau\omega = \omega^s$ pour un certain s, et

$$\sigma\tau\omega = \sigma(\omega^s) = \sigma(\omega)^s = \omega^{rs} = \tau\sigma\omega.$$

Par suite $\sigma\tau = \tau\sigma$, et le groupe de Galois est abélien, comme il fallait le montrer.

(2) *Soit K un corps; supposons que les racines n-ièmes de l'unité soient dans K. Soit $a \in K$. Soit α une racine du polynôme $t^n - a$, de sorte que $\alpha^n = a$, et soit $L = K(\alpha)$. Nous allons encore montrer que L est abélien sur K.* Soit σ un K-plongement de L. Alors

$$(\sigma\alpha)^n = \sigma(\alpha^n) = \sigma a = a.$$

Par suite, $\sigma\alpha$ est aussi racine de $t^n - a$, et

$$(\alpha/\sigma\alpha)^n = 1.$$

Par suite, si ω est une racine primitive de l'unité, il y a un entier r tel que

$$\sigma\alpha = \omega^r\alpha.$$

En particulier, L est galoisienne sur K. Si τ est un K-automorphisme de L, alors, pour un entier s

$$\tau\alpha = \omega^s\alpha$$

d'où

$$\sigma\tau\alpha = \sigma(\omega^s\alpha) = \omega^s\sigma\alpha = \omega^s\omega^r\alpha = \tau\sigma\alpha.$$

Par conséquent, $\sigma\tau = \tau\alpha$ et le groupe de Galois est encore abélien.

Soit K un corps et f un polynôme de degré $\geqq 1$. Nous dirons que f est *résoluble*

par radicaux, si son corps de dislocation est contenu dans une extension galoisienne L qui admet une suite

$$K = K_0 \subset K_1 \subset K_2 \subset \cdots \subset K_m = L$$

de sous-corps de L, telle que

(a) $K_1 = K(\omega)$, pour une racine primitive n-ième ω de l'unité,

(b) pour tout i tel que $1 \leqq i \leqq m - 1$, le corps K_{i+1} peut s'exprimer sous la forme $K_{i+1} = K_i(\alpha_{i+1})$, où α_{i+1} est une racine d'un certain polynôme

$$t^d - a_i = 0,$$

où d divise n, et a_i est élément de K_i.

Remarquons que si d divise n, alors $\omega^{n/d}$ est une racine primitive d-ième de l'unité (pourquoi?) et par suite, d'après ce que nous avons vu, l'extension K_{i+1} de K_i est abélienne. Nous avons vu également que K_1 est abélien sur K. L'extension K est donc décomposée en une suite d'extensions abéliennes. Soit G_i le groupe de Galois de L sur K_i. On obtient alors une suite correspondante de sous-groupes

$$G \supset G_1 \supset G_2 \supset \cdots \supset G_m = \{e\}$$

tel que G_{i+1} soit distingué dans G_i, et le groupe quotient G_i/G_{i+1} abélien (cf. exercice 4 du §4). On a démontré, par conséquent, que

***Théorème* 13.** *Si f est résoluble par radicaux, alors son groupe de Galois est résoluble.*

C'était un problème célèbre que de déterminer si tout polynôme était résoluble par radicaux (*Galois, jeune mathématicien révolutionnaire, le résolut* - N.d.T.). Pour montrer que ce n'est pas le cas, il va suffire d'exhiber un polynôme dont le groupe de Galois est le groupe symétrique S_5 (ou S_n, pour $n \geqq 5$), d'après le théorème 8 du chapitre II, §5. Cela est facile:

***Théorème* 14.** *Soient $x_1, \ldots, x_n$ des éléments algébriquement indépendants sur un corps K, et soit*

$$f(t) = \prod_{i=1}^{n} (t - x_i) = t^n - s_1 t^{n-1} + \cdots + (-1)^n s_n$$

où

$$s_1 = x_1 + \cdots + x_n, \ldots, s_n = x_1 \cdots x_n$$

sont les coefficients de f. Soit $K = K_0(s_1, \ldots, s_n)$. Soit $L = K(x_1, \ldots, x_n)$. Alors L est galoisienne sur K, de groupe de Galois S_n.

Démonstration. L'extension L est galoisienne sur K puisque

$$L = K(x_1, \ldots, x_n)$$

est le corps de dislocation de f. Etant donnée une permutation σ de $\{1, \ldots, n\}$, on sait qu'il existe un automorphisme π_σ du corps $K_0(x_1, \ldots, x_n)$ laissant K_0 fixé, et tel que $\pi_\sigma(x_i) = x_{\sigma(i)}$. Puisque une permutation quelconque de $x_1, \ldots, x_n$ laisse les coefficients de f invariants, il s'ensuit que K est fixe sous chaque π_σ pour chaque

$\sigma \in S_n$, et nous voyons que l'application $\sigma \longrightarrow \pi\sigma$ fournit un homomorphisme injectif de S_n dans le groupe de Galois de L sur K. Cependant, nous avons également constaté qu'on peut représenter tout K-automorphisme de L comme permutation des racines de f. Il s'ensuit que S_n représente tout automorphisme du groupe de Galois, comme il fallait le montrer.

Dans le paragraphe suivant, nous montrons qu'on peut toujours trouver n nombres complexes algébriquement indépendants sur $\mathbf{Q}$.

§7. *Extensions infinies*

Commençons par quelques assertions de cardinalité concernant les corps. Nous n'utilisons que des ensembles finis ou dénombrables ici, et tout ce que nous avons besoin de savoir sur de tels ensembles est ce qui suit:

Si D est dénombrable, un produit fini $D \times \cdots \times D$ est dénombrable.
Une réunion dénombrable d'ensembles dénombrables est dénombrable,
Un sous-ensemble infini d'un ensemble dénombrable est dénombrable,
Si D est dénombrable et si $D \to E$ est une application surjective de D sur un ensemble E qui n'est pas fini, alors E est dénombrable.

Le lecteur trouvera des démonstrations indépendantes de ces assertions au chapitre VIII (cf. théorème 3 et ses corollaires), mais celles-ci ne sont que de simples exercices.

Soit K un corps et E une extension de K. Nous disons que E est algébrique sur K si tout élément de E est algébrique sur K. Soit $\bar{K}$ l'ensemble de tous les nombres complexes algébriques sur K. Alors $\bar{K}$ est un corps, car nous avons vu que si α et β sont algébriques sur K, alors $\alpha + \beta$ et $\alpha\beta$ le sont aussi, puisque contenus dans l'extension finie $K(\alpha, \beta)$ de K.

***Théorème* 15.** *Soit K un ensemble dénombrable. Alors $\bar{K}$ est dénombrable.*

Démonstration. Procédons par étapes. Soit P_n l'ensemble des polynômes irréductibles de degré $n \geqq 1$, à coefficients dans K et de coefficients dominants égaux à 1. A chaque polynôme $f \in P_n$,

$$f(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0,$$

nous associons ses coefficients $(a_{n-1}, \ldots, a_0)$. On obtient ainsi une injection de P_n dans $K \times \cdots \times K = K^n$, d'où nous concluons que P_n est dénombrable.

Soient maintenant $\alpha_{f,1}, \ldots, \alpha_{f,n}$ les racines d'un polynôme $f \in P_n$, dans un ordre donné.

Soit $J_n = \{1, \ldots, n\}$, et soit

$$P_n \times \{1, \ldots, n\} \to \mathbf{C}$$

l'application de $P_n \times J_n$ dans $\mathbf{C}$ telle que

$$(f, i) \mapsto \alpha_{f,i}$$

pour $i = 1, \ldots, n$ et $f \in P_n$. Cette application est une surjection de $P_n \times J_n$ sur l'ensemble des nombres de degré n sur K, et, par suite, cet ensemble est dénombrable. En prenant la réunion pour $n = 1, 2, \ldots$, on arrive à la conclusion que l'ensemble des nombres algébriques sur K est dénombrable.

***Théorème* 16.** *Soit K un ensemble dénombrable. Le corps $K(t)$ des fractions rationnelles est alors dénombrable.*

Démonstration. Il va suffire de démontrer que l'anneau des polynômes $K[t]$ est dénombrable, parce que nous avons une application surjective

$$K[t] \times K[t]_0 \to K(t),$$

où $K[t]_0$ désigne l'ensemble des éléments non nuls de $K[t]$. L'application est naturellement $(a, b) \mapsto a/b$. Pour chaque n, soit P_n l'ensemble des polynômes de degré $\leqq n$, à coefficients dans K. Alors P_n est dénombrable, et, par suite, $K[t]$ est dénombrable, comme étant réunion dénombrable de P_0, P_1, $P_2, \ldots$, avec $\{0\}$.

Corollaire. *Etant donné un entier $n \geqq 1$, il existe n nombres complexes algébriquement indépendants sur* $\mathbf{Q}$.

Démonstration. Le corps $\overline{\mathbf{Q}}$ est dénombrable, et $\mathbf{C}$ ne l'est pas. Par suite, il existe $x_1 \in \mathbf{C}$ qui est transcendant sur $\overline{\mathbf{Q}}$. Soit $K_1 = \overline{\mathbf{Q}}(x_1)$. Alors K_1 est dénombrable. En procédant par récurrence, on peut prendre un élément x_2 transcendant sur $\overline{K}_1$, et ainsi de suite, pour trouver les éléments $x_1, \ldots, x_n$ que nous souhaitons.

Remarque. Le fait que $\mathbf{C}$ (et même $\mathbf{R}$) n'est pas dénombrable sera démontré au chapitre suivant.

CHAPITRE VII

Les nombres réels et les nombres complexes

§1. *Anneaux ordonnés*

Soit A un anneau intègre. On entend par *ordre* sur l'anneau A un sous-ensemble P de A satisfaisant les conditions suivantes:

ORD 1. *Pour tout $x \in A$ on a soit $x \in P$, soit $x = 0$, soit $-x \in P$, et chacune de ces trois possibilités est exclusive des deux autres.*

ORD 2. *Si $x, y \in P$, alors $x + y \in P$ et $xy \in P$.*

On dit aussi que A est un *anneau ordonné* par P, et on appelle P l'ensemble des éléments *positifs*.

Supposons que A est ordonné par P. Puisque $1 \neq 0$, et puisque $1 = 1^2 = (-1)^2$, on voit que 1 est un élément de P, i.e. que 1 est positif. D'après ORD 2, et par récurrence, il s'ensuit que $1 + \cdots + 1$ (n fois) est positif. Un élément $x \in A$ tel que $x \neq 0$ et $x \notin P$ est dit *négatif.* Si x et y sont des éléments négatifs de A, alors xy est positif (parce que $-x \in P$, $-y \in P$ et donc $(-x)(-y) = xy \in P$). Si x est positif et y négatif, alors xy est négatif, parce que $-y$ est positif, et que, par conséquent, $x(-y) = -xy$ est positif. Pour tout $x \in A$, $x \neq 0$, on voit que x^2 est positif.

Supposons que A est un corps. Si x est positif et si $x \neq 0$, alors $xx^{-1} = 1$, et il résulte des remarques précédentes que x^{-1} aussi est positif.

Soit A un anneau ordonné quelconque, et soit A' un sous-anneau de A. Soit P l'ensemble des éléments positifs de A et soit $P' = P \cap A'$. Il est alors clair que P' définit un ordre sur A', qu'on appelle l'ordre induit sur A'.

Soient, plus généralement, A' et A des anneaux ordonnés et P' et P leurs ensembles respectifs d'éléments positifs. Soit $f : A' \to A$ un plongement (i.e. un homomorphisme injectif). Nous disons que f *respecte l'ordre* si, pour tout $x \in A'$ tel que $x > 0$, nous avons $f(x) > 0$. Cela revient à dire que $f^{-1}(P) = P'$ (ou $f^{-1}(P)$ est l'ensemble de tous les $x \in A'$ tels que $f(x) \in (P)$.

Soit $x, y \in A$. On pose $x < y$ (ou $y > x$) pour dire que $y - x \in P$. Dire, dans ces conditions, que $x < 0$ revient à dire que x est négatif, ou que $-x$ est positif. On vérifie facilement les inégalités habituelles, à savoir, que pour $x, y, z \in A$,

IN 1. $x < y$ et $y < z$ impliquent $x < z$,
IN 2. $x < y$ et $z > 0$ impliquent $xz < yz$,
IN 3. $x < y$ implique $x + z < y + z$,

si A est un corps, alors

IN 4. $x < y$ et $x, y > 0$ impliquent $1/y < 1/x$.

Démontrons IN 2, à titre d'exemple. On a $y - x \in P$ et $z \in P$, de sorte que, d'après ORD 2, $(y - x)z \in P$. Mais $(y - x)z = yz - xz$, de sorte que, par définition, $xz < yz$. Démontrons IN 4, pour donner un autre exemple: on multiplie l'inégalité $x < y$ par x^{-1} et y^{-1} et on trouve l'assertion IN 4. Les autres assertions ci-dessus sont laissées en exercices.

Si $x, y \in A$, on pose $x \leqq y$ pour dire que $x < y$ ou que $x = y$. On vérifie alors immédiatement que IN 1, IN 2 et IN 3 restent valables si on y remplace partout le signe $<$ par $\leqq$. En outre, on vérifie tout aussi immédiatement que si $x \leqq y$ et si $y \leqq x$, alors $x = y$.

On va voir au théorème suivant comment on peut étendre l'ordre sur un anneau intègre à un ordre de son corps de fractions.

***Théorème* 1.** *Soit A un anneau intègre, ordonné par P. Soit K son corps de fractions. Soit P_K l'ensemble des éléments de K qu'on peut écrire sous la forme a/b, avec $a, b \in A$, $b > 0$ et $a > 0$. Alors P_K définit un ordre sur K.*

Démonstration. Soit $x \in K, x \neq 0$. On peut écrire, en multipliant numérateur et dénominateur de x par -1, s'il le faut, l'élément x sous la forme a/b, avec $a, b \in A$ et $b > 0$. Si $a > 0$, alors $x \in P_K$. Si $-a > 0$, alors $-x = -a/b \in P_K$. On ne peut pas avoir à la fois x et $-x \in P_K$, car sinon, on pourrait écrire

$$x = a/b \qquad \text{et} \qquad -x = c/d,$$

avec $a, b, c, d \in A$ et $a, b, c, d > 0$. Alors on

$$-a/b = c/d,$$

d'où $-ad = bc$. Mais $bc \in P$ et $ad \in P$, ce qui est contradictoire. Cela prouve que P_K satisfait à ORD 1. Soient maintenant $x, y \in P_K$; posons

$$x = a/b \qquad \text{et} \qquad y = c/d,$$

avec $a, b, c, d \in A$ et $a, b, c, d > 0$. On a alors $xy = ac/bd \in P_K$. La fraction

$$x + y = \frac{ad + bc}{bd}$$

est aussi dans P_K. Cela prouve que P_K satisfait à ORD 2, et démontre notre théorème.

Le théorème montre en particulier comment étendre l'ordre habituel de l'anneau **Z** des entiers au corps **Q** des rationnels. La façon de définir les entiers rationnels et de les ordonner, est étudiée en appendice.

Soit A un anneau ordonné comme précédemment. Si $x \in A$, on pose

$$|x| = \begin{cases} x \text{ si } x \leqq 0, \\ -x \text{ si } x > 0. \end{cases}$$

Nous avons alors la caractérisation de la fonction $x \mapsto |x|$, qu'on appelle la *valeur absolue:*

Pour chaque $x \in A$, $|x|$ *est l'unique* $z \in A$ *tel que* $z \geqq 0$ *et* $z^2 = x^2$.

Pour démontrer cela, remarquons d'abord qu'on a certainement $|x|^2 = x^2$, et $|x| \geqq 0$, pour tout $x \in A$. D'autre part, étant donné $a \in A, a > 0$, il existe au plus deux éléments $z \in A$ tels que $z^2 = a$, parce que le polynôme $t^2 - a$ possède au plus deux racines. Si $w^2 = a$, alors $w \neq 0$ et $(-w)^2 = w^2 = a$. Il y a, par conséquent, au plus un élément positif $z \in A$ tel que $z^2 = a$. Cela démontre notre assertion.

Définissons le symbole $\sqrt{a}$ pour $a \geqq 0$ dans A comme étant l'élément $z \geqq 0$ de A tel que $z^2 = a$, lorsqu'un tel z existe. Sinon, $\sqrt{a}$ n'est pas défini. Il est maintenant facile de voir que, si $a, b \geqq 0$ et si $\sqrt{a}$ et $\sqrt{b}$ existent, alors $\sqrt{ab}$ existe et

$$\sqrt{ab} = \sqrt{a}\sqrt{b}.$$

En effet, si $z, w \geqq 0$ et $z^2 = a$, $w^2 = b$, alors $(zw)^2 = z^2w^2 = ab$. On peut alors exprimer la définition de la valeur absolue au moyen de l'expression $|x| = \sqrt{x^2}$.

La valeur absolue vérifie les règles suivantes:

VA 1. *Pour tout* $x \in A$, *on a* $|x| \geqq 0$ *et* $|x| > 0$ *si* $x \neq 0$,
VA 2. $|xy| = |x||y|$, *pour tous* $x, y \in A$,
VA 3. $|x + y| \leqq |x| + |y|$, *pour tous* $x, y \in A$.

La première assertion ci-dessus est évidente. Quant à VA 2, on a

$$|xy| = \sqrt{(xy)^2} = \sqrt{x^2y^2} = \sqrt{x^2}\sqrt{y^2} = |x||y|.$$

On a, en ce qui concerne VA 3,

$$\begin{aligned} |x + y|^2 = (x + y)^2 &= x^2 + xy + xy + y^2 \\ &\leqq |x|^2 + 2|xy| + |y|^2 \\ &= |x|^2 + 2|x||y| + |y|^2 \\ &= (|x| + |y|)^2. \end{aligned}$$

On peut alors en déduire le résultat en prenant les racines carrées des deux membres de l'égalité précédente (nous avons utilisé ici deux propriétés des inégalités, cf. exercice 1).

EXERCICES

1. Soit A un anneau intègre ordonné. (a) Démontrez que $x \leqq |x|$, pour tout $x \in A$. (b) Si $a, b \geqq 0$, si $a \leqq b$, et si $\sqrt{a}$ et $\sqrt{b}$ existent, montrez que $\sqrt{a} \leqq \sqrt{b}$.

2. Soit K un corps ordonné. Soit P l'ensemble des polynômes

$$f(t) = a_nt^n + \cdots + a_0$$

à coefficients dans K, où $a_n > 0$. Montrez que P définit un ordre sur $K[t]$.

3. Soit A un anneau intègre ordonné. Si $x, y \in A$, montrez que $|-x| = x$,

$$|x - y| \geqq |x| - |y|,$$

et aussi que

$$|x + y| \geqq |x| - |y|.$$

Démontrez également que $|x| \leqq |x + y| + |y|$.

4. Soit K un corps ordonné et $f : \mathbf{Q} \to K$ un plongement du corps des rationnels dans K. Montrez que f respecte nécessairement l'ordre.

§2. *Préliminaires*

Soit K un corps ordonné. De l'exercice 4 du paragraphe précédent, on tire que le plongement de $\mathbf{Q}$ dans K, respecte l'ordre. Nous identifions $\mathbf{Q}$ à un sous-corps de K.

Rappelons une définition formelle. Soit E un ensemble. Une *suite* d'éléments de E est simplement une application.

$$\mathbf{Z}^+ \to E$$

de l'ensemble des entiers positifs dans E. On désigne habituellement une suite par la notation

$$\{x_1, x_2, \ldots\}$$

ou par

$$\{x_n\}_{n \geqq 1}$$

ou simplement par

$$\{x_n\},$$

s'il n'y a aucun risque de confusion avec l'ensemble réduit à x_n.

Une suite $\{x_n\}$ de K est dite *suite de Cauchy* si, étant donné un élément $\varepsilon > 0$ de K, il existe un entier positif N tel que pour tous les entiers $m, n \geqq N$, on ait

$$|x_n - x_m| \leqq \varepsilon.$$

(Pour plus de simplicité, nous faisons la convention que N, n, m désignent des entiers positifs à moins d'autres précisions. Convenons également de ce que ε désigne des éléments de K.)

Pour éviter l'usage de trop de symboles, nous disons qu'un énoncé donné A concernant des entiers positifs est vrai pour tous les entiers *suffisamment ou assez grands*, s'il existe N tel que l'énoncé $S(n)$ est vrai pour tout $n \geqq N$. Il est clair que, si $S_1, \ldots, S_r$ sont des énoncés, tous vrais pour des entiers suffisamment grands, alors ces énoncés sont simultanément vérifiés pour des entiers suffisamment grands. En effet, si

$$S_1(n) \text{ est vrai pour } n \geqq N_1, \ldots, S_r(n) \text{ est vrai pour } n \geqq N_r,$$

on voit, en posant $N = \max(N_1, \ldots, N_r)$, que $S_i(n)$ est vérifié pour $n \geqq N$.

Nous disons qu'un énoncé est vrai pour des entiers *arbitrairement grands* si, étant donné N, l'énoncé est vérifié pour un $n \geqq N$.

Une suite $\{x_n\}$ de K est dite *convergente* s'il existe un élément $x \in K$ tel que, étant donné $\varepsilon > 0$, on a

$$|x - x_n| \leqq \varepsilon,$$

pour tous les n assez grands.

Un corps ordonné dans lequel toute suite de Cauchy est convergente est dit *complet*. Remarquons que le nombre x ci-dessus, s'il existe, est déterminé de façon unique car si $y \in K$ est tel que

$$|y - x_n| \leqq \varepsilon$$

pour des n assez grands, alors

$$|x - y| \leqq |x - x_n + x_n - y| \leqq |x - x_n| + |x_n - y| \leqq 2\varepsilon.$$

Comme cela est vrai pour tout $\varepsilon > 0$ de K, il s'ensuit que $x - y = 0$, et donc que $x = y$. On appelle ce nombre de x, la *limite de la suite* $\{x_n\}$.

On dira qu'un corps K est *archimédien* si, étant donné $x \in K$, il existe un entier positif n tel que $x \leqq n$. Il résulte alors, qu'étant donné $\varepsilon > 0$ de K, on peut trouver un entier $m > 0$ tel que $1/\varepsilon < m$, donc tel que $1/m < \varepsilon$.

Il est facile de dire que le corps des rationnels n'est pas complet. On peut, par exemple construire des suites de Cauchy de rationnels dont le carré approche 2, mais tels que la suite n'a pas de limite dans $\mathbf{Q}$, (sinon $\sqrt{2}$ serait rationnel). Au paragraphe suivant, nous allons construire un corps archimédien complet, qui est appelé le corps des nombres réels. Nous allons maintenant démontrer une propriété des corps archimédiens, qu'on prend comme point de départ de l'analyse.

Soit E un sous-ensemble de K. Par *majorant* de E, on tend un élément $z \in K$ tel que $x \leqq z$, pour tout $x \in E$. Par *borne supérieure* de E, on entend un élément $w \in K$ tel que w soit un majorant, et tel que, si z est un majorant, alors $w \leqq z$. Si w_1, w_2 sont deux bornes supérieures, alors $w_1 \leqq w_2$ et $w_2 \leqq w_1$, de sorte que $w_1 = w_2$: une borne supérieure est unique.

Théorème 2. *Soit K un corps ordonné archimédien complet. Alors toute partie non vide E de K qui possède un majorant, a aussi une borne supérieure.*

Démonstration. Considérons, pour tout entier positif n, l'ensemble T_n constitué de tous les entiers y tels que, pour tout $x \in E$, nous avons $nx \leqq y$ (et par suite $x \leqq y/n$). L'ensemble T_n est donc minoré par tout élément nx (où $x \in E$), et n'est pas vide car si b est un majorant de E, alors tout entier y tel que $nb \leqq y$ est dans T_n (en utilisant la propriété qu'a le corps d'être archimédien). Soit y_n le plus petit élément de T_n. Il existe alors un élément x_n de E tel que

$$y_n - 1 < nx_n \leqq y_n$$

(sinon, y_n n'est pas le plus petit élément de T_n). Par suite,

$$\frac{y_n}{n} - \frac{1}{n} < x_n \leqq \frac{y_n}{n}.$$

Soit $z_n = y_n/n$. Nous affirmons que la suite $\{z_n\}$ est de Cauchy. Pour le démontrer, considérons deux entiers positifs m et n; supposons par exemple que $y_n/n \leqq y_m/m$. Nous affirmons que

$$\frac{y_m}{m} - \frac{1}{m} < \frac{y_n}{n} \leqq \frac{y_m}{m}.$$

En effet, sinon

$$\frac{y_n}{n} \leqq \frac{y_m}{m} - \frac{1}{m}$$

et

$$\frac{y_m}{m} - \frac{1}{m}$$

est un majorant de E, ce qui n'est pas vrai puisque x_m lui est supérieur. Cela démontre notre affirmation, à partir de laquelle nous constatons que

$$\left|\frac{y_n}{n} - \frac{y_m}{m}\right| \leqq \frac{1}{m}.$$

Pour m et n assez grands, $1/m$ est arbitrairement petit, et nous avons démontré que notre suite $\{z_n\}$ est de Cauchy.

Soit w sa limite. Démontrons d'abord que w majore E. Supposons qu'il existe $x \in E$ tel que $w < x$. Il existe alors un n tel que

$$|z_n - w| \leqq \frac{x - w}{2}.$$

Alors

$$\begin{aligned} x - z_n = x - w + w - z_n &\geqq x - w - |w - z_n| \\ &\geqq x - w - \frac{x - w}{2} \\ &\geqq \frac{x - w}{2} > 0. \end{aligned}$$

de sorte que $x > z_n$, contrairement au fait que z_n est un majorant de E.

Nous allons maintenant montrer que w est une borne supérieure de E. Soit $u < w$. Il existe un n tel que

$$|z_n - x_n| \leqq \frac{1}{n} < \frac{w - u}{4}.$$

(Il faut simplement choisir n assez grand.) On peut également choisir n assez grand pour que

$$|z_n - w| \leqq \frac{w-u}{4},$$

puisque w est la limite de z_n. Mais alors

$$\begin{aligned} x_n - u &= w - u + x_n - z_n + z_n - w \\ &\geqq w - u - |x_n - z_n| - |z_n - w| \\ &\geqq w - u - \frac{w-u}{4} - \frac{w-u}{4} \\ &\geqq \frac{w-u}{2} > 0, \end{aligned}$$

d'où $u < x_n$. Par suite u n'est pas la borne supérieure de E. Cela prouve que w est la borne supérieure de E, et achève la preuve de notre théorème.

§3. *Construction de nombres réels*

Partons de l'ensemble **Q** des rationnels et de leur ordre obtenu à partir de l'ordre de l'ensemble des entiers du théorème 1 du §1. Nous voulons définir les nombres réels. Dans l'enseignement élémentaire, on se sert des réels sous forme de nombres décimaux illimités, comme

$$\sqrt{2} = 1,414...$$

Un tel nombre décimal illimité n'est rien d'autre qu'une suite de rationnels

$$1 \;; \quad 1,4 \;; \quad 1,41 \;; \quad 1,414 \;;$$

et il est bon de remarquer qu'il existe d'autres suites qui «approchent» $\sqrt{2}$. Si l'on veut *définir* $\sqrt{2}$, il est alors raisonnable de dire que c'est, par définition, une suite de rationnels, pour une notion convenable d'équivalence. Nous allons faire cela pour tous les nombres réels.

Partons de notre corps ordonné **Q** et des suites de Cauchy de **Q**. Soit $\gamma = \{c_n\}$ une suite de nombres rationnels. On dira que γ est une *suite nulle*, (i.e. *tendant vers* 0. N.d.T.) si, étant donné un nombre rationnel $\epsilon > 0$, nous avons

$$|c_n| \leqq \epsilon$$

pour des n assez grands. A moins d'autres précisions, nous avons affaire dans ce qui suit à des rationnels, et nos suites sont des suites de nombres rationnels.

Si $\alpha = \{a_n\}$ et $\beta = \{b_n\}$ sont des suites de nombres rationnels, on définit $\alpha + \beta$ comme étant la suite $\{a_n + b_n\}$, i.e. la suite dont le n-ième terme est $a_n + b_n$. On définit la suite $\alpha\beta$ comme la suite dont le n-ième terme est $a_n b_n$. L'ensemble des suites de rationnels n'est donc rien d'autre que l'anneau de toutes les applications

de $\mathbf{Z}^+$ dans $\mathbf{Q}$. Nous allons voir dans un instant que les suites de Cauchy forment un sous-anneau de cet anneau.

Lemme 1. *Soit $\alpha = \{a_n\}$ une suite de Cauchy. Il existe un nombre rationnel positif B tel que $|a_n| \leqq B$, pour tout n.*

Démonstration. Etant donné le nombre 1, il existe N tel que pour $n \geqq N$, nous ayons

$$|a_n - a_N| \leqq 1.$$

On a alors, pour tout $n \geqq N$,

$$|a_n| \leqq |a_N| + 1.$$

Soit B le maximum de $|a_1|, |a_2|, \ldots, |a_{N-1}|, |a_N| + 1$.

Lemme 2. *Les suites de Cauchy forment un anneau commutatif.*

Démonstration. Soient $\alpha = \{a_n\}$ et $\beta = \{b_n\}$ des suites de Cauchy. Etant donné $\epsilon > 0$, nous avons

$$|a_n - a_m| \leqq \frac{\epsilon}{2}$$

pour tout m et tout n assez grands, ainsi que

$$|b_n - b_m| \leqq \frac{\epsilon}{2}$$

pour tous n et m suffisamment grands. Par suite, pour tous n et m assez grands, nous avons

$$\begin{aligned} |a_n + b_n - (a_m + b_m| &= |a_n - a_m + b_n - b_m| \\ &\leqq |a_n - a_m| + |b_n - b_m| \\ &\leqq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

La somme $\alpha + \beta$ est donc de Cauchy. On voit immédiatement que

$$-\alpha = \{-a_n\}$$

est une suite de Cauchy. Quant au produit, on a

$$\begin{aligned} |a_n b_n - a_m b_m| &= |a_n b_n - a_n b_m + a_n b_m - a_m b_m| \\ &\leqq |a_n|\,|b_n - b_m| + |a_n - a_m|\,|b_m|. \end{aligned}$$

Il existe, d'après le lemme 1, $B_1 > 0$ tel que $|a_n| \leqq B_1$ pour tout n, et $B_2 > 0$ tel que $|b_n| \leqq B_2$, pour tout n. Soit $B = \max(B_1, B_2)$. Pour tous m et n assez grands, nous avons

$$|a_n - a_m| \leqq \frac{\epsilon}{2B} \qquad \text{et} \qquad |b_n - b_m| \leqq \frac{\epsilon}{2B},$$

et par conséquent,

$$|a_n b_n - a_m b_m| \leqq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Le produit $\alpha\beta$ est ainsi une suite de Cauchy. Il est clair que la suite $\{1, 1, 1, \ldots\}$ est de Cauchy. Par conséquent, les suites de Cauchy forment un sous-anneau de toutes les applications de $\mathbf{Z}^+$ dans $\mathbf{Q}$. Cet anneau est évidemment commutatif.

Lemme 3. *Les suites nulles forment un idéal de l'anneau des suites de Cauchy.*

Démonstration. Soient $\beta = \{b_n\}$ et $\gamma = \{c_n\}$ des suites nulles. Etant donné $\varepsilon > 0$ nous avons, pour des n assez grands

$$|b_n| \leqq \frac{\epsilon}{2} \qquad \text{et} \qquad |c_n| \leqq \frac{\epsilon}{2}.$$

Ainsi, pour tout n suffisamment grand, a

$$|b_n + c_n| \leqq \epsilon,$$

de telle sorte que $\beta + \gamma$ est une suite nulle. Il est clair que $-\beta$ est une suite nulle.

D'après le lemme 1, étant donné une suite de Cauchy $\alpha = \{a_n\}$, il existe un nombre rationnel $B > 0$ tel que

$$|a_n| \leqq B$$

pour tout n. Pour tout n assez grand, on a

$$|b_n| \leqq \frac{\epsilon}{B},$$

d'où

$$|a_n b_n| \leqq B \frac{\epsilon}{B} = \epsilon,$$

de telle sorte que $\alpha\beta$ est une suite nulle. Cela prouve, comme souhaité, que les suites nulles forment un idéal.

Soit A l'anneau des suites de Cauchy et M l'idéal des suites nulles. Nous avons donc ici une notion d'équivalence, qui est, par définition que, si $\alpha, \beta \in A$, $\alpha \equiv \beta \pmod{M}$ signifie $\alpha - \beta \in M$, ou, autrement dit que $\alpha = \beta + \gamma$, pour une suite γ nulle. Par définition, un *nombre réel* est une classe d'équivalence de suites de Cauchy modulo M. Comme nous l'avons vu à propos de la construction des anneaux quotients, l'ensemble de telles classes d'équivalence est lui-même un anneau, désigné par A/M, mais que nous notons aussi $\mathbf{R}$. La classe d'équivalence de α est pour l'instant notée $\bar{\alpha}$. Par définition, on a alors

$$\overline{\alpha + \beta} = \bar{\alpha} + \bar{\beta}, \; \overline{\alpha\beta} = \bar{\alpha}\bar{\beta}.$$

L'élément unité de $\mathbf{R}$ est la classe de la suite de Cauchy $\{1, 1, 1, \ldots\}$.

Théorème 3. *L'anneau $A/M = \mathbf{R}$ des nombres réels est en fait un corps.*

Démonstration. On doit démontrer que, si α est une suite de Cauchy, et n'est pas la suite nulle, alors il existe une suite de Cauchy telle que $\alpha\beta \equiv e \pmod{M}$, où $e = \{1, 1, 1, \ldots\}$. Nous avons besoin d'un lemme concernant les suites nulles.

Lemme 4. *Soit α une suite de Cauchy non nulle. Il existe alors N_0 et un rationnel $c > 0$ tels que $|a_n| \geqq c$, pour tout $n \geqq N_0$.*

Démonstration. Supposons le contraire. Soit $\alpha = \{a_n\}$. Alors, étant donné $\varepsilon > 0$, il existe une suite infinie $n_1 < n_2 < \ldots$ d'entiers positifs tels que

$$|a_{n_i}| < \frac{\epsilon}{3},$$

pour tout $i = 1, 2, \ldots$ Par définition, il existe N tel que pour $m, n \geqq N$, nous avons

$$|a_n - a_m| \leqq \frac{\epsilon}{3}.$$

Soit $n_i \geqq N$. Nous avons, pour $m \geqq N$,

$$|a_m| \leqq |a_m - a_{n_i}| + |a_{n_i}| \leqq \frac{2\epsilon}{3},$$

et pour $m, n \geqq N$,

$$|a_n| \leqq |a_m| + \frac{\epsilon}{3} \leqq \epsilon.$$

Cela montre que α est une suite nulle, contrairement à l'hypothèse, et démontre notre lemme.

Revenons à la démonstration du théorème. D'après le lemme 4, il existe N_0 tel que, pour $n \geqq N_0$, nous avons $a_n \neq 0$. Soit $\beta = \{b_n\}$ la suite telle que $b_n = 1$ si $n < N_0$ et $b_n = a_n^{-1}$ si $n < N_0$. Alors $\beta\alpha$ ne diffère de e que par un nombre fini de termes, de sorte que $\beta\alpha - e$ est certainement une suite nulle. Reste à démontrer que β est une suite de Cauchy. D'après le lemme 4, on peut choisir N_0 tel que, pour tout $n \geqq N_0$, nous avons $a_n \geqq c > 0$. Il s'ensuit que

$$\frac{1}{|a_n|} \leqq \frac{1}{c}.$$

Etant donné $\epsilon > 0$, il existe N (que l'on peut prendre $\geqq N_0$) tel que, pour tout $m, n \geqq N$, nous ayons

$$|a_n - a_m| \leqq \epsilon c^2.$$

Nous obtenons alors, pour $m, n \geqq N$

$$\left|\frac{1}{a_n} - \frac{1}{a_m}\right| = \left|\frac{a_m - a_n}{a_m a_n}\right| \leqq \frac{\epsilon c^2}{c^2} = \epsilon,$$

prouvant par là que β est une suite de Cauchy et concluant la démonstration de notre théorème.

Nous avons construit le corps des nombres réels.

Remarquons que nous avons un homomorphisme d'anneaux de **Q** dans **R**, en envoyant chaque nombre rationnel a sur la classe de la suite de Cauchy $\{a, a, a, \ldots\}$. Cet homomorphisme est le composé de deux homomorphismes: l'application

$$a \mapsto \{a, a, a, \ldots\}$$

de **Q** dans l'anneau des suites de Cauchy suivie de l'application $A \to A/M$. Puisqu'il ne s'agit pas de l'homomorphisme nul, on tire un isomorphisme de **Q** sur son image.

Le lemme qui suit a pour but de définir un ordre sur les nombres réels.

Lemme 5. *Soit $\alpha = \{a_n\}$ une suite de Cauchy. Une et une seule des assertions suivantes esr alors vérifiée:*

(1) *α est une suite nulle.*

(2) *il existe un rationnel $c > 0$ tel que, pour tous les n assez grands, $a_n \geqq c$,*

(3) *il existe un rationnel $c < 0$ tel que, pour tous les n assez grands, $a_n \leqq c$.*

Démonstration. Il est clair que si α vérifie une des trois assertions, α ne vérifie pas les deux autres, i.e. que ces assertions s'excluent mutuellement. Ce que nous devons montrer, c'est qu'au moins l'une des trois possibilités est réalisée. Supposons que α n'est pas une suite nulle. D'après le lemme 4, il existe N_0 et un nombre rationnel $c > 0$ tel que $|a_n| \geqq c$ pour tout $n \geqq N_0$. On a donc $a_n \geqq c$ si a_n est positif, et $-a_n \geqq c$ si a_n est négatif. Supposons qu'il existe des entiers n assez grands pour que a_n soit négatif. Supposons qu'il existe des entiers n assez grands pour que a_n soit positif, et des entiers m assez grands pour que a_m soit négatif. Alors pour de tels m et n, nous avons

$$a_n - a_m \geqq 2c > 0,$$

ce qui contredit le fait que α est une suite de Cauchy. Cela prouve que (2) ou (3) doivent être vérifiées, et achève la démonstration du lemme.

Lemme 6. *Soit $\alpha = \{a_n\}$ une suite de Cauchy et soit $\beta = \{b_n\}$ une suite nulle. Si α vérifie la propriété (2) du lemme 5, alors $\alpha + \beta$ aussi et, si α vérifie la propriété (3) du lemme 5, alors $\alpha + \beta$ aussi.*

Démonstration. Supposons que α satisfait à la propriété (2). Pour tout n assez grand et par définition d'une suite nulle, on a $|b_n| \leqq c/2$. Par suite, pour des n assez grands,

$$a_n + b_n \geqq |a_n| - |b_n| \geqq c/2.$$

Un raisonnement similaire démontre le résultat analogue pour la propriété (3). Cela démontre le lemme.

On peut maintenant définir un ordre sur les nombres réels. Soit P l'ensemble des réels qui sont représentés par une suite α de Cauchy ayant la propriété (2); démontrons que P définit un ordre.

Soit α une suite de Cauchy représentant un nombre réel. Si α n'est pas nulle et ne vérifie pas (2), alors $-\alpha$ satisfait évidement (2). D'après le lemme 6, toute autre suite de Cauchy représentant le même nombre réel que α satisfait aussi (2). Par suite P satisfait à la condition ORD 1.

Soient $\alpha = \{a_n\}$ et $\beta = \{b_n\}$ deux suites de Cauchy représentant des nombres réels de P et vérifiant (2). Il existe $c_1 > 0$ tel que $a_n \geqq c_1$ pour n assez grand et il existe $c_2 > 0$ tel que $b_n \geqq c_2$ pour n assez grand. Par suite $a_n + b_n \geqq c_1 + c_2 > 0$ pour n assez grand, prouvant ainsi que $\alpha + \beta$ est aussi dans P. De plus,

$$a_n b_n \geqq c_1 c_2 > 0,$$

pour n suffisamment grand, de telle sorte que $\alpha\beta$ est dans P. Cela démontre que P définit un ordre sur le corps des nombres réels.

Rappelons-nous que nous avons trouvé un isomorphisme de Q sur un sous-corps de **R**, isomorphisme donné par l'application

$$a \mapsto \{\overline{a, a, a, \ldots}\}.$$

En vertu de l'exercice 4, §1, cette application respecte l'ordre, mais cela résulte tout aussi facilement de nos définitions. Pendant un moment, nous n'identifions pas a avec son image dans **R**, et nous désignons par $\bar{a}$ la classe de la suite de Cauchy $\{a, a, a, \ldots\}$.

***Théorème* 4.** *L'ordre de* **R** *est archimédien.*

Démonstration. Soit A un nombre réel, représenté par une suite de Cauchy $\alpha = \{a_n\}$. D'après le lemme 1, on peut trouver un nombre rationnel r tel que $a_n \leqq r$ pour tout n, et en multipliant par un dénominateur positif, on voit qu'il existe un entier b tel que $a_n \leqq b$, pour tout n. La classe $\bar{b} - \bar{\alpha}$ est alors réprésentée par la suite $\{b - a_n\}$ et $b - a_n \geqq 0$, pour tout n. Il résulte que, par définition,

$$\bar{b} - \bar{\alpha} \geqq 0,$$

d'où $\bar{\alpha} \leqq \bar{b}$, comme on le veut.

Les lemmes suivants nous donnent un critère d'inégalités entre nombres réels en termes de suite de Cauchy.

***Lemme* 7.** *Soit $\gamma = \{c_n\}$ une suite de Cauchy de nombres rationnels, et soit c un nombre rationnel > 0. Si $|c_n| \geqq c$ pour des n assez grands, alors $|\bar{\gamma}| \leqq \bar{c}$.*

Démonstration. Si $\bar{\gamma} = 0$, notre assertion est triviale. Supposons $\gamma \neq 0$, et soit par exemple $\bar{\gamma} > 0$. On a alors $|\bar{\gamma}| = \bar{\gamma}$, et nous devons donc montrer que $\bar{c} - \bar{\gamma} \geqq 0$. Mais, pour n assez grand,

$$c - c_n \geqq 0.$$

Puisque $\bar{c} - \bar{\gamma} = \{\bar{c} - \bar{c}_n\}$, il résulte de notre définition de l'ordre dans A que $\bar{c} - \bar{\gamma} \geqq 0$. Le cas où $\bar{\gamma} < 0$ se démontre en considérant $-\bar{\gamma}$.

Etant donné un nombre $\epsilon > 0$, d'après le théorème 4, il existe un nombre rationnel $\epsilon_1 > 0$ tel que $0 < \epsilon_1 < \epsilon$. Par suite de la définition de limite, le fait de prendre ϵ donné réel ou rationnel est sans importance.

Lemme 8. *Soit $\alpha = \{\alpha_n\}$ une suite de Cauchy de rationnels. La classe α est la limite de la suite $\{a_n\}$.*

Démonstration. Etant donné un nombre rationnel $\epsilon > 0$, il existe N tel que, pour $m, n \geqq N$, nous ayons

$$|a_n - a_m| \leqq \epsilon.$$

Nous avons alors, d'après le lemme 7, pour tous $m, n \geqq N$

$$|\alpha - \bar{a}_m| = |\{\overline{a_n - a_m}\}| \leqq \epsilon.$$

Cela prouve notre assertion.

Théorème 5. *Le corps des nombres réels est complet.*

Démonstration. Soit $\{A_n\}$ une suite de Cauchy de nombres réels. On peut trouver pour tout n, d'après le lemme 8, un nombre rationnel a_n tel que

$$|A_n - a_n| \leqq \tfrac{1}{n}.$$

(On doit, pour être rigoureux, écrire $1/n$ dans le membre de droite!) De plus, par définition, étant donné $\epsilon > 0$, il existe N tel que, pour tous $m, n \geqq N$, nous avons

$$|A_n - A_m| \leqq \frac{\epsilon}{3}.$$

Soit N_1 un entier $\geqq N$, et tel que $1/N_1 \leqq \epsilon/3$. Alors, pour tous $m, n \geqq N_1$, nous obtenons

$$\begin{aligned}|\bar{a}_n - \bar{a}_m| &= |\bar{a}_n - A_n + A_n - A_m + A_m - \bar{a}_m| \\ &\leqq |\bar{a}_n - A_n| + |A_n - A_m| + |A_m - \bar{a}_m| \\ &\leqq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.\end{aligned}$$

Cela prouve que $\{\bar{a}_n\}$ est une suite de Cauchy de nombres rationnels. Soit A sa limite. On a, pour tout n

$$|A_n - A| \leqq |A_n - \bar{a}_n| + |\bar{a}_n - A|.$$

On voit, en prenant n assez grand, que A est aussi limite de la suite $\{A_n\}$, démontrant ainsi notre théorème.

Exercices

1. Soit p un nombre premier. Si x est un nombre rationnel non nul, écrit sous la forme $x = p^r a/b$ où r est un entier, a et b des entiers non divisibles par p, on pose

$$|x|_p = 1/p^r.$$

et $|0|_p = 0$. Montrez que, pour tous rationnels x, y, on a

$$|xy|_p = |x|_p |y|_p \qquad \text{et} \qquad |x + y|_p \leqq |x|_p + |y|_p.$$

2. Définissez des suites de Cauchy respectant p, et construisez une complétion de $\mathbf{Q}$ respectant cette fonction, qu'on appelle *valeur absolue p-adique.*

3. Démontrez que tout nombre réel positif possède une racine carrée positive dans $\mathbf{R}$. Puisque le polynôme $t^2 - a$ a au plus deux racines dans un corps, et puisque, pour toute racine α, le nombre $-\alpha$ est aussi racine, montrez que pour tout $a \in \mathbf{R}$, $a > 0$, il existe un unique $\alpha \in \mathbf{R}$, $\alpha \geqq 0$ tel que $\alpha^2 = a$. [*Indication:* soit α la borne supérieure de l'ensemble des nombres rationnels b tels que $b^2 \leqq a$.]

§4. *Développements décimaux*

Théorème 6. *Soit d un entier $\geqq 2$, et soit m un entier $\geqq 0$. On peut alors écrire m d'une façon unique sous la forme*

$$m = c_0 + c_1 d + \cdots + c_n d^n$$

pour des entiers c_i tels que $0 \leqq c_i < d$.

Démonstration. On le voit facilement, par division euclidienne, et nous allons donner cette démonstration. Posons pour démontrer l'existence de cette écriture $c_0 = m$ et $c_i = 0$ pour $i > 0$, si $m < d$. Si $m < d$ on écrit

$$m = qd + c_0$$

avec $0 \leqq c_0 < d$, en utilisant la division euclidienne. On a alors $q < m$ et, par hypothèse de récurrence, il existe des entiers $c_i(0 \leqq c_i < d$ et $i \geqq 1)$ tels que

$$q = c_1 + c_2 d + \cdots + c_k d^k.$$

En remplaçant par cette valeur, on obtient ce qu'on veut. Quant à l'unicité, supposons que

$$m = b_0 + b_1 d + \cdots + b_n d^n$$

avec des entiers b_1 satisfaisant $0 \leqq b_i < d$. (On peut utiliser le même n, à condition d'ajouter des termes ayant des coefficients b_i ou c_i nuls, s'il le faut.) Supposons, par exemple, que $a_0 \leqq b_0$. Alors $b_0 - a_0 \geqq 0$ et $b_0 - a_0 < d$. D'autre part $b_0 - a_0 = d$ pour un entier e [ce qu'on voit en soustrayant (2) de (1)]. Par suite $b_0 - a_0 = 0$ et $b_0 = a_0$. Supposons que nous avons montré que $a_i = b_i$ pour $0 \leqq i \leqq s$ et $s < n$. Alors

$$a_{s+1} d^{s+1} + \cdots + a_n d^n = b_{s+1} d^{s+1} + \cdots + b_n d_n.$$

En divisant les deux côtés par d^{s+1}, on obtient

$$a_{s+1} + \cdots + a_{n-s-1}d^{n-s-1} = b_{s+1} + \cdots + b_{n-s-1}d^{n-s-1}.$$

De ce que nous venons de voir, il résulte que $a_{s+1} = b_{s+1}$, et nous avons prouvé l'unicité souhaitée, par récurrence.

Soit x un nombre positif réel, et d un entier $\geqq 2$. L'élément x possède une unique expression de la forme

$$x = m + \alpha$$

où $0 \leqq \alpha < 1$. Soit en effet m le plus grand entier $\leqq x$. Alors $x < m + 1$, et par suite $0 \leqq x - m < 1$. Nous allons maintenant expliciter un développement d-décimal en base d pour les nombres réels compris entre 0 et 1.

Théorème 7. *Soit x un nombre réel tel que $0 \leqq x < 1$. Soit d un entier $\geqq 2$. Il y a pour chaque entier positif une expression unique*

$$x = \frac{a_1}{d} + \frac{a_2}{d^2} + \cdots + \frac{a_n}{d^n} + \alpha_n \tag{3}$$

pour des entiers a_i satisfaisant à $0 \leqq a_i < d$ et $0 \leqq \alpha_n < 1/d^n$.

Démonstration. Soit m le plus grand entier $\leqq d^n x$. Alors $m \geqq 0$ et

$$d^m x = m + \alpha_n$$

pour des nombres α_n tels que $0 \leqq \alpha_n < 1$. Appliquons le théorème 6 à m, et divisons ensuite par d^n pour obtenir l'expression désirée. Réciproquement, étant donnée une telle expression (3), on peut la multiplier par d^n et appliquer le résultat d'unicité contenu dans le théorème 6 pour en tirer l'unicité de (3). Cela démontre notre théorème.

Lorsque $d = 10$, les nombres $a_1, a_2, \ldots$ du théorème 7 sont précisément ceux du développement décimal de x, qui s'écrit

$$x = 0, a_1 a_2 a_3 \ldots,$$

depuis des temps immémoriaux.

Réciproquement,

Théorème 8. *Soit d un entier $\geqq 2$. Soient $a_1, a_2, \ldots$ une suite d'entiers tels que $0 \leqq a_i < d$, pour tout i; supposons qu'étant donné un entier positif N, il existe un $n \geqq N$ tel que $a_n \neq d - 1$. Il existe alors un nombre réel x tel que, pour tout $n \geqq 1$, on a*

$$x = \frac{a_1}{d} + \frac{a_2}{d^2} + \cdots + \frac{a_n}{d^n} + \alpha_n$$

où α_n est un nombre tel que $0 \leqq \alpha_n < 1/d^n$.

Démonstration. Nous prenons la liberté d'utiliser quelques propriétés élémen-

taires des limites et des sommes infinies, développées dans tous les exposés d'initiation à l'analyse. Soit

$$y_n = \frac{a_1}{d} + \cdots + \frac{a_n}{d^n}.$$

La suite $y_1, y_2, \ldots$ est donc croissante et on montre facilement qu'elle est majorée. Soit x sa borne supérieure. L'élément x est alors limite de la suite, et

$$x = y_n + \alpha_n,$$

où

$$\alpha_n = \sum_{\nu=n+1}^{\infty} \frac{a_\nu}{d^\nu},$$

soit

$$\beta_n = \sum_{\nu=n+1}^{\infty} \frac{d-1}{d^\nu}.$$

On a, par hypothèse, $\alpha_n < \beta_n$ puisque il existe un a_ν tel que $\nu \geqq n+1$ pour lequel $a_\nu \neq d-1$. D'autre part

$$\beta_n = \frac{d-1}{d^{n+1}} \sum_{\nu=0}^{\infty} \frac{1}{d^\nu} = \frac{d-1}{d^{n+1}} \frac{1}{1-\frac{1}{d}} = \frac{1}{d^n}.$$

D'où $0 \leqq \alpha_n < 1/d^n$, comme il fallait le démontrer.

Corollaire. *L'ensemble des nombres réels n'est pas dénombrable.*

Démonstration. Considérons le sous-ensemble de l'ensemble des nombres constitués des suites décimales

$$0, a_1 a_2 \ldots$$

avec $0 \leqq a_i \leqq 8$, en faisant $d = 10$ dans les théorèmes 7 et 8. Il va suffire de démontrer que ce sous ensemble n'est pas dénombrable. Supposons le contraire, et soit

$$\begin{aligned} \alpha_1 &= 0, a_{11} a_{12} a_{13} \ldots, \\ \alpha_2 &= 0, a_{21} a_{22} a_{23} \ldots, \\ \alpha_3 &= 0, a_{31} a_{32} a_{33} \ldots, \\ &\ldots \end{aligned}$$

une énumération de ce sous-ensemble. Soit $b_1, b_2, \ldots$ une suite d'entiers positifs telle que $0 \leqq b_i \leqq 8$ et $b_i \neq a_{ii}$ pour tout i. Soit $\alpha = 0, b_1 b_2 b_3 \ldots$

Cet α n'est égal à aucun des $\alpha_n (n = 1, 2, \ldots)$. Cela contredit l'hypothèse stipulant que l'on a une énumération de ce sous-ensemble, et démontre ainsi notre corollaire. (*Remarque:* les résultats élémentaires concernant la théorie des ensembles dénombrables utilisés dans cette démonstration sont développés systématiquement dans le prochain chapitre.)

Exercices

Jetez un oeil sur un texte traitant d'approximations diophantiennes ou de fractions continues, pour voir comment on définit, au moyen de la division euclidienne, une suite plus naturelle (canonique) de nombres rationnels convergeant vers un nombre réel x donné, et indépendante du système décimal (qui est un système particulier).

§5. *Les nombres complexes*

Notre but dans ce paragraphe est d'identifier l'ensemble des nombres réels avec un sous-corps d'un corps dans lequel l'équation $t^2 = -1$ a une racine. Comme d'habitude en l'occurrence, on définit ce plus grand corps de façon à rendre évidente la résolution de cette équation, et nous devons ensuite démontrer les propriétés requises.

Définissons un *nombre complexe* comme étant un couple (x, y) de nombres réels. Si $z = (x, y)$, on définit la multiplication par un nombre réel a comme étant donnée par

$$az = (ax, ay).$$

A ce stade, l'ensemble des nombres complexes n'est donc rien d'autre que $\mathbf{R}^2$, et on peut déjà le considérer comme espace vectoriel sur $\mathbf{R}$. Posons $e = (1, 0)$ et $i = (0, 1)$. Tout nombre complexe peut donc s'exprimer d'une façon unique comme $xe + yi$, pour x et y réels. Nous devons maintenant définir la multiplication des nombres complexes. Si $z = xe + yi$ et $w = ue + vi$ sont deux nombres complexes, avec $x, y, u, v \in \mathbf{R}$, on pose

$$zw = (xu - yv)e + (xv + yu)i.$$

Remarquons tout de suite que $ez = ze = z$, pour tout $z \in \mathbf{C}$, et que $i^2 = -e$. Nous affirmons que $\mathbf{C}$ est un corps. Nous savons déjà que c'est un groupe (abélien) additif. Si $z_1 = x_1 e + y_1 i$, $z_2 = x_2 e + y_2 i$ et $z_3 = x_3 e + y_3 i$, alors

$$\begin{aligned}(z_1 z_2)(z_3) &= \big((x_1 x_2 - y_1 y_2)e + (y_1 x_2 + x_1 y_2)i\big)(x_3 e + y_3 i)\\ &= (x_1 x_2 x_3 - y_1 y_2 x_3 - y_1 x_2 y_3 - x_1 y_2 y_3)e\\ &\quad + (y_1 x_2 x_3 + x_1 y_2 x_3 + x_1 x_2 y_3 - y_1 y_2 y_3)i.\end{aligned}$$

Un calcul analogue de $z_1(z_2z_3)$ montre qu'on obtient là même valeur que pour $(z_1z_2)z_3$. De plus, en posant $w = u + vi$ une fois encore, on a

$$\begin{aligned} w(z_1 + z_2) &= (ue + vi)\big((x_1 + x_2)e + (y_1 + y_2)i\big) \\ &= \big(u(x_1 + x_2) - v(y_1 + y_2)\big)e + \big(v(x_1 + x_2) + u(y_1 + y_2)\big)i \\ &= (ux_1 - vy_1 + ux_2 - vy_2)e + (vx_1 + uy_1 + vx_2 + uy_2)i. \end{aligned}$$

Un calcul direct de $wz_1 + wz_2$ montre qu'on obtient la même chose que pour $w(z_1 + z_2)$. Nous avons évidemment aussi $wz = zw$, pour tous $w, z \in \mathbf{C}$, et par suite, $(z_1 + z_2)w = z_1w + z_2w$. Cela prouve que les nombres complexes constituent un anneau commutatif.

On vérifie immédiatement que l'application $x \mapsto (x, 0)$ est un homomorphisme injectif de $\mathbf{R}$ dans $\mathbf{C}$, et à partir de maintenant, nous identifions $\mathbf{R}$ et son image dans $\mathbf{C}$, c'est-à-dire que nous écrivons x au lieu de xe, pour tout $x \in \mathbf{R}$.

Si $z = x + iy$ est un nombre complexe, on définit son *conjugué* $\bar{z}$ par

$$\bar{z} = x - iy.$$

On voit, d'après les règles de multiplication, que

$$z\bar{z} = x^2 + y^2.$$

Si $z \neq 0$, l'un au moins des nombres x ou y n'est pas nul, et on voit que

$$\lambda = \frac{\bar{z}}{x^2 + y^2}$$

est tel que $z\lambda = \lambda z = e$, parce que

$$z\frac{\bar{z}}{x^2 + y^2} = \frac{z\bar{z}}{x^2 + y^2} = 1$$

Par suite tout élément non nul de $\mathbf{C}$ a un inverse, et par conséquent, $\mathbf{C}$ est un corps, qui contient $\mathbf{R}$ comme sous-corps [modulo notre identification de x avec $(x, 0)$].

On définit le *module* d'un nombre complexe $z = x + iy$ comme étant

$$|z| = \sqrt{a^2 + b^2}$$

et en termes de module, on peut écrire l'inverse d'un nombre complexe non nul sous la forme

$$z^{-1} = \frac{\bar{z}}{|z|^2}.$$

Soient z et w deux nombres complexes; on voit facilement que

$$|z + w| \leqq |z| + |w| \qquad \text{et} \qquad |zw| = |z||w|.$$

De plus, $\overline{z + w} = \bar{z} + \bar{w}$ et $\overline{zw} = \bar{z}\bar{w}$. Nous laissons la démonstration en exercices. Nous avons maintenant amené la théorie des nombres complexes au point où les analystes la prennent en charge.

En particulier, si un nombre complexe $z = x + iy$ a pour valeur absolue 1,

alors $x^2 + y^2 = 1$, et à partir de là, il existe un nombre réel θ tel que $x = \cos\theta$ et $y = \sin\theta$. Nous avons la définition $e^{i\theta} = \cos\theta + \sin\theta$, et pour tout nombre complexe z, nous avons la définition $e^z = e^x e^{iy}$. Des formules additives du sinus et du cosinus, nous concluons alors que $e^{z+w} = e^z e^w$ pour tout nombre complexe z, w. Par ailleurs, si $z \neq 0$, alors $z/|z|$ a une valeur absolue égale à 1, et de là tout nombre complexe z peut être écrit sous la forme polaire $z = re^{i\theta}$ où r est un nombre réel $\geqq 0$ et θ un nombre réel. En utilisant la fonction réelle exponentielle, on prouve que tout nombre réel positif r possède une racine $n^{ième}$ réelle, et en utilisant la forme polaire, on conclut que tout nombre complexe a une racine $n^{ième}$ complexe, à savoir $r^{1/n}e^{i\theta/n}$.

En dehors de ce fait, nous utilisons la propriété qu'a une fonction à valeurs réelles, sur un ensemble fermé borné de nombres complexes, d'avoir un maximum. Tout cela est démontré dans les traités élémentaires d'analyse.

En utilisant ces résultats, nous allons maintenant démontrer que *le corps des nombres complexes est algébriquement clos, ou, en d'autres termes, que tout polynôme* $f \in \mathbf{C}[t]$, *de degré* $\geqq 1$, *possède une racine dans* $\mathbf{C}$.

On peut écrire

$$f(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_0,$$

avec $a_n \neq 0$. Pour tout réel $R > 0$, la fonction $|f|$ telle que $t \mapsto |f(t)|$ est continue sur le disque fermé de rayon R et possède, par suite, un minimum sur ce disque. D'autre part, de l'expression

$$f(t) = a_n t^n \left(1 + \frac{a_{n-1}}{a_n t} + \cdots + \frac{a_0}{a_n t^n}\right),$$

on tire que, lorsque $|t|$ devient grand, $|f(t)|$ aussi, i.e. qu'étant donné $C > 0$, il existe $R > 0$ tel que si $|t| > R$, alors $|f(t)| > C$. Par conséquent, il existe un nombre positif R_0 tel que, si z_0 donne un minimum local de $|f|$ sur le disque fermé de rayon R_0, alors

$$|f(t)| \geqq |f(z_0)|,$$

pour tout nombre complexe t. Autrement dit, z_0 donne un minimum absolu de $|f|$. Nous allons montrer que $f(z_0) = 0$.

Exprimons f sous la forme

$$f(t) = c_0 + c_1(t - z_0) + \cdots + c_n(t - z_0)^n,$$

pour des constantes c_i. Si $f(z_0) \neq 0$, alors $c_0 = f(z_0) \neq 0$. Soit $z = t - z_0$ et soit m le plus petit entier > 0 tel que $c_m \neq 0$. Cet entier m existe parce que f est supposé de degré $\geqq 1$. On peut alors écrire

$$f(t) = f_1(z) = c_0 + c_m z^m + z^{m+1} g(z)$$

pour des polynômes g, et pour un polynôme f_1 (obtenu à partir de f par changement de variable). Soit z_1 un nombre complexe tel que $z_1^m = -c_0/c_m$, et considérons des valeurs de z du type

$$z = \lambda z_1,$$

où λ est réel tel que $0 \leqq \lambda \leqq 1$. On a

$$\begin{aligned} f(t) = f_1(\lambda z_1) &= c_0 - \lambda^m c_0 + \lambda^{m+1} z_1^{m+1} g(\lambda z_1) \\ &= c_0[1 - \lambda^m + \lambda^{m+1} z_1^{m+1} c_0^{-1} g(\lambda z_1)]. \end{aligned}$$

Il existe un nombre $C > 0$ tel que pour tout λ vérifiant $0 \leqq \lambda \leqq 1$, nous avons $|z_1^{m+1} c_0^{-1} g(\lambda z_1)| \leqq C$ et, par suite,

$$|f_1(\lambda z_1)| \leqq |c_0|(1 - \lambda^m + C\lambda^{m+1}).$$

Si l'on peut maintenant démontrer que, pour λ assez petit vérifiant $0 < \lambda < 1$, nous avons

$$0 < 1 - \lambda^m + C\lambda^{m+1} < 1.$$

nous obtenons $|f_1(\lambda z_1)| < |c_0|$ pour de tels λ, ce qui contredit l'hypothèse $|f(z_0)| \leqq |f(t)|$ pour tout nombre complexe t. L'inégalité à gauche est évidente puisque $0 < \lambda < 1$. L'inégalité à droite équivaut à $C\lambda^{m+1} < \lambda^m$, ou encore à $C\lambda < 1$, certainement vérifié pour des λ assez petits. Cela achève la démonstration.

Remarque. L'idée de la démonstration est en fait assez simple. Nous avons notre polynôme

$$f_1(z) = c_0 + c_m z^m + z^{m+1} g(z),$$

avec $c_m \neq 0$. Si $g = 0$, nous ne faisons qu'amener $c_m z^m$ à l'origine en soustrayant un terme de même direction que c_0. On peut le faire en extrayant la racine n-ième convenable comme ci-dessus. Puisque g est en général non nul, nous avons à effectuer quelques manipulations analytiques pour montrer que le troisième terme est très petit, par rapport à $c_m z^m$, et qu'il ne trouble pas le cours général de la démonstration de façon essentielle.

Exercices

1. En supposant connu le résultat qu'on vient de montrer, démontrer que tout polynôme irréductible sur le corps des nombres réels est de degré 1 ou 2. [*Indication:* décomposer le polynôme sur le corps des complexes et prendre les couples de racines complexes conjuguées.]

2. Démontrez qu'un polynôme irréductible de degré 2 sur **R**, de coefficient dominant égal à 1, peut s'écrire

$$(t - a)^2 + b^2$$

avec $a, b \in \mathbf{R}$ et $b > 0$.

CHAPITRE VIII

Ensembles

§1. *Un peu de vocabulaire*

Ce chapitre est le plus abstrait du livre, et c'est le seul qui traite d'objets ayant de si pauvres structures, à savoir justement des ensembles. Le point remarquable est que l'on peut, avec si peu de choses en main, démontrer des résultats intéressants.

Nous allons d'abord définir quelques termes. Soient E et I des ensembles. On entend par *famille d'éléments de E, indexée par I* tout simplement une application $f : I \to E$. Cependant, quand nous parlons d'une famille, nous écrivons plutôt f_i que $f(i)$, et nous utilisons aussi la notation $\{f_i\}_{i \in I}$ pour désigner la famille.

Exemple 1. Soit E l'ensemble constitué du seul élément 3. Soit $I = \{1, ..., n\}$ l'ensemble des entiers de 1 à n. Une famille d'éléments de E, indexée par I, peut s'écrire $\{a_i\}_{i=1,\cdots,n}$ où tous les a_i sont égaux à 3. Remarquons qu'une famille n'est pas la même chose qu'un sous-ensemble. Le même élément de E peut recevoir des indices distincts.

Une famille d'éléments de E indexée par des entiers positifs, ou non négatifs, est aussi dite une *suite*.

Exemple 2. On écrit souvent une suite de nombres réels sous la forme

$$\{x_1, x_2, ...\} \qquad \text{ou} \qquad \{x_n\}_{n \geqq 1}$$

qui remplace l'application $f : \mathbf{Z}^+ \to \mathbf{R}$ telle que $f(i) = x_i$. Comme précédemment remarquons qu'une suite peut avoir tous ses éléments égaux, comme

$$\{1, 1, 1, ...\}$$

qui est une suite d'entiers, où $x_i = 1$ pour tout $i \in \mathbf{Z}^+$.

On définit une *famille d'ensembles indexée par un ensemble I* de la même façon, c'est-à-dire qu'une famille d'ensembles indexée par I est une correspondance

$$i \mapsto E_i$$

qui à chaque $i \in I$ associe un ensemble E_i. Les ensembles E_i peuvent avoir ou ne pas avoir d'éléments en commun, et il est concevable qu'ils soient égaux. On écrit, comme précédemment $\{E_i\}_{i \in I}$ la famille.

On peut définir l'intersection et la réunion de familles d'ensembles, exactement comme l'intersection et la réunion d'un nombre fini d'ensembles. Ainsi, si $\{E_i\}_{i\in I}$ est une famille d'ensembles, on définit l'*intersection* de cette famille comme étant l'ensemble

$$\bigcap_{i\in I} E_i$$

constitué de tous les éléments x qui sont dans tous les E_i. On définit la *réunion*

$$\bigcup_{i\in I} E_i$$

comme étant l'ensemble de tous les x appartenant à un E_i au moins.

Si E et E' sont des ensembles, on définit $E \times E'$ comme l'ensemble de tous les couples (x, y) où $x \in E$ et $y \in E'$. On peut définir des produits finis analogues de la même façon. Si $E_1, E_2, \ldots$ est une suite d'ensembles, on définit le produit

$$\prod_{i=1}^{\infty} E_i$$

comme étant l'ensemble de toutes les suites $(x_1, x_2, \ldots)$, où $x_i \in E_i$. De façon analogue, si I est un ensemble d'indices, et $\{E_i\}_{i\in I}$ une famille d'ensembles, on définit le produit

$$\prod_{i\in I} E_i$$

comme étant l'ensemble de toutes les familles $\{x_i\}_{i\in I}$, où $x_i \in E_i$.

Soient X, Y et Z des ensembles. On a la formule

$$(X \cup Y) \times Z = (X \times Z) \cup (Y \times Z).$$

Pour le démontrer, considérons $(w, z) \in (X \cup Y) \times Z$, où $w \in X \cup Y$ et où $z \in Z$. On a donc $w \in X$ ou $w \in Y$. Soit par exemple $w \in X$. Alors $(w, z) \in X \times Z$. Donc

$$(X \cup Y) \times Z \subset (X \times Z) \cup (Y \times Z).$$

Réciproquement, $X \times Z$ est contenu dans $(X \cup Y) \times Z$ et $Y \times Z$ aussi. Par conséquent leur réunion est contenue dans $(X \cup Y) \times Z$, prouvant par là notre assertion.

On dit que deux ensembles X et Y sont *disjoints* si leur intersection est vide. On dit que la réunion $X \cup Y$ est *disjointe* si X et Y sont disjoints. Remarquons que si X et Y sont disjoints, $(X \times Z)$ et $(Y \times Z)$ le sont aussi. On peut prendre des produits pour des familles quelconques. Par exemple si $\{X_i\}_{i\in I}$ est une famille d'ensembles, alors

$$\Big(\bigcup_{i\in I} X_i\Big) \times Z = \bigcup_{i\in I} (X_i \times Z).$$

Si la famille $\{X_i\}_{i\in I}$ est disjointe (c'est-à-dire si $X_i \cap X_j$ est vide pour $i \neq j$ lorsque $i,j \in I$), alors les ensembles $X_i \times Z$ sont également disjoints.

On a des formules analogues pour l'intersection. Par exemple.

$$(X \cap Y) \times Z = (X \cap Z) \times (Y \cap Z).$$

Nous en laissons la démonstration au lecteur.

Soit X un ensemble et Y une partie de X. Le *complémentaire* de Y dans X, désigné par $\complement_X Y$, ou par $X - Y$, est l'ensemble des éléments $x \in X$ tels que $x \notin Y$. Si Y et Z sont des parties de X, on a les formules suivantes

$$\complement_X(Y \cup Z) = \complement_X Y \cap \complement_X Z,$$

$$\complement_X(Y \cap Z) = \complement_X Y \cup \complement_X Z.$$

Ces dernières ne sont que des formulations de définitions. Supposons, par exemple, que $x \in X$ et que $x \notin (Y \cup Z)$. Alors $x \notin Y$ et $x \notin Z$. Donc $x \in \complement_X Y \cap \complement_X Z$. Réciproquement si $x \in \complement_X Y \cap \complement_X Z$, alors x n'est ni dans Y, ni dans Z, et par suite $x\ \complement_X(Y \cup Z)$. Cela démontre la première formule et nous laissons la seconde au lecteur.

Exercice. Explicitez de telles formules pour le *complémentaire* de l'union d'une famille d'ensembles, et pour le complémentaire de l'intersection d'une famille d'ensembles.

Soient A et B des ensembles et $f : A \to B$ une application. Si Y est une partie de B, on définit $f^{-1}(Y)$ comme l'ensemble de tous les $x \in A$ tels que $f(x) \in Y$. Il se peut que $f^{-1}(Y)$ soit vide, naturellement. On appelle $f^{-1}(Y)$ l'*image réciproque* de Y (par f). Si f est injective et si Y est réduite à exactement un élément, alors $f^{-1}(y)$ est soit vide, soit réduite à exactement un élément. Nous allons donner en exercices certaines propriétés simples de l'image réciproque.

EXERCICES

1. Soient $f : A \to B$ une application, Y et Z des sous-ensembles de B; établissez les formules suivantes:

$$f^{-1}(Y \cup Z) = f^{-1}(Y) \cup f^{-1}(Z),$$

$$f^{-1}(Y \cap Z) = f^{-1}(Y) \cap f^{-1}(Z).$$

2. Formulez et démontrez les propriétés analogues à celles de l'exercice 1, pour des familles de parties; montrez, par exemple, que si $\{Y_i\}_{i\in I}$ est une famille de parties de B,

$$f^{-1}\Big(\bigcup_{i\in I} Y_i\Big) = \bigcup_{i\in I} f^{-1}(Y_i).$$

3. Soit $f : A \to B$ une application surjective. Montrez qu'il existe une application injective de B dans A.

§2. *Lemme de Zorn*

Dans le but de manipuler efficacement une infinité d'ensembles simultanément, nous avons besoin d'un axiome particulier. Introduisons un peu de vocabulaire pour l'énoncer.

Soit E un ensemble. Un *ordre partiel* (aussi appelé un ordre) de E est une relation, écrite $x \leqq y$, entre certains couples d'éléments de E, ayant les propriétés suivantes:

OP. 1. *on a* $x \leqq x$;
OP. 2. *si* $x \leqq y$ *et* $y \leqq z$, *alors* $x \leqq z$;
OP. 3. *si* $x \leqq y$ *et* $y \leqq x$, *alors* $x = y$.

Remarquons que nous n'exigeons pas que la relation $x \leqq y$ ou $y \leqq x$ soit vérifiée par tous les couples (x, y) d'éléments de E. Certains couples peuvent ne pas être comparables. On écrit parfois $y \geqq x$ pour $x \leqq y$.

Exemple 1. Soit G un groupe. Soit E l'ensemble de ses sous-groupes. Si H et H' sont des sous-groupes de G, on pose

$$H \leqq H'$$

si H est un sous-groupe de H'. On vérifie immédiatement que cette relation définit un ordre partiel sur E. Etant donnés deux sous-groupes H et H' de G, on n'a pas nécessairement $H \leqq H'$ ou $H' \leqq H$.

Exemple 2. Soit A un anneau, et soit E l'ensemble des idéaux à gauche de A. On définit un ordre partiel sur E d'une façon analogue à la précédente, c'est-à-dire en posant, pour des idéaux à gauche I et I' de A,

$$I \leqq I'$$

si $I \subset I'$.

Exemple 3. Soit X un ensemble et E l'ensemble des parties de X. Si Y et Z sont des parties de X, on pose $Y \leqq Z$, si Y est un sous-ensemble de E. Cela définit un ordre partiel sur E.

Dans tous ces exemples, la relation d'ordre partiel définie est dite relation d'inclusion.

Si $x \leqq y$ et $x \neq y$ dans un ensemble partiellement ordonné, on écrit alors $x < y$.

Remarque. Nous n'avons pas défini le mot relation, mais cela peut se faire en termes d'ensembles comme suit. Par définition, une *relation* entre couples d'éléments d'un ensemble A est une partie R du produit $A \times A$. Si $x, y \in A$ et $(x, y) \in R$, alors on dit que x et y *vérifient notre relation.* En utilisant cette formulation, nous pouvons rétablir nos conditions d'ordre partiel sous la forme qui suit: pour tous $x, y, z \in A$,

OP 1. *On a* $(x, x) \in R$.
OP 2. *Si* $(x, y) \in R$ *et* $(y, z) \in R$, *alors* $(x, z) \in R$.
OP 3. *Si* $(x, y) \in R$ *et* $(y, x) \in R$, *alors* $x = y$.

La notation précédemment utilisée est cependant beaucoup plus facile à manier, et, après avoir montré comment cette notation peut s'expliciter en termes d'ensembles, nous continuons à utiliser la première définition que nous avons donnée.

Soient A un ensemble partiellement ordonné, et B une partie de A. On peut alors définir un ordre partiel sur B en posant $x \leqq y$ pour $x, y \in B$ si et seulement si $x \leqq y$ dans A. Autrement dit, si $R \subset A \times A$ est la partie de $A \times A$ définissant notre relation d'ordre partiel sur A, nous posons $R_0 = R \cap (B \times B)$, et R_0 définit une relation d'ordre partiel sur B. On dit que R_0 est l'*ordre partiel induit* sur B par R, ou bien est la *restriction* à B de l'ordre partiel de A.

Soit E un ensemble partiellement ordonné. Par *minimum* (ou *plus petit élément*) de E, on entend un élément $a \in E$ tel que $a \leqq x$ pour tout $x \in E$. De la même façon par *plus grand élément* (ou *maximum*, N.d.T.), on entend un élément b tel que $x \leqq b$, pour tout $x \in E$.

On entend par *élément maximal* m de E un élément tel que si $x \in E$ et $x \geqq m$, alors $x = m$. Remarquons qu'un élément maximal n'est pas nécessairement un plus grand élément. Il peut y avoir beaucoup d'éléments maximaux dans E, tandis que s'il existe un plus grand élément, il est unique (démonstration ?).

Soit E un ensemble partiellement ordonné. On dit que E est *totalement ordonné* si, étant donnés x et y dans E, on a nécessairement $x \leqq y$ ou $y \leqq x$.

Exemple 4. Les entiers de $\mathbf{Z}$ sont totalement ordonnés par l'ordre usuel. Ainsi en est-il aussi des nombres réels.

Soient E un ensemble partiellement ordonné et T une partie de E. Un *majorant* de T (dans E) est un élément $b \in E$ tel que $x \leqq b$, pour tout $x \in T$. Une *borne supérieure* de T dans E est un majorant b tel que, si c est un autre majorant, alors $b \leqq c$. On dira qu'un ensemble ordonné E est *inductif* si toute partie non vide totalement ordonnée possède un majorant.

On dira que l'ensemble ordonné E est *strictement inductif* si toute partie non vide *de E* possède une borne supérieure.

Dans les exemples 1, 2, 3, l'ensemble ordonné est, à chaque fois, strictement inductifs. Plaçons-nous, par exemple, dans le cas de l'exemple 1 pour le montrer. Soit T une partie non vide totalement ordonnée de l'ensemble des sous-groupes de G. Cela veut dire que si H et H' sont dans T alors $H \subset H'$ ou $H' \subset H$. Soit U la réunion de tous les ensembles de T. Alors:

(1) U est un sous-groupe. *Démonstration:* si $x, y \in U$, il existe des sous-groupes H et H' de T tels que $x \in H$ et $y \in H'$. Soit, par exemple, $H \subset H'$ et alors les deux éléments x et y sont dans H' et par suite $xy \in H'$. Par suite $xy \in U$. On a aussi $x^{-1} \in H'$ et $x^{-1} \in U$. Par conséquent U est un sous-groupe.

(2) U est un majorant de tout élément de T. *Démonstration:* tout $H \in T$ est contenu dans U, de sorte que $H \leqq U$, pour tout $H \in T$.

(3) U est une borne supérieure de T. *Démonstration:* tout sous-groupe de G qui contient tous les sous-groupes de T contient leur réunion.

La démonstration du fait que les ensembles des exemples 2 et 3 sont strictement inductifs est entièrement analogue.

Nous allons maintenant énoncer l'axiome mentionné au début du paragraphe.

Lemme de Zorn. *Il existe dans tout ensemble ordonné non vide et inductif un élément maximal.*

Nous allons voir deux exemples de la façon d'appliquer le lemme de Zorn.

Théorème 1. *Soit A un anneau non commutatif d'élément unité $1 \neq 0$. Il existe dans A un idéal maximal.*

(Rappelons qu'un idéal maximal est un idéal M tel que $M \neq A$, et tel que, si J est un idéal vérifiant $M \subset J \subset A$, alors $J = M$ ou $J = A$.)

Démonstration. Soit E l'ensemble des idéaux propres de A, qui sont des idéaux tels que $J \neq A$. L'ensemble E n'est donc pas vide, puisque l'idéal nul est dans E. De plus, E est inductif pour l'inclusion. Pour le voir, considérons une partie non vide totalement ordonnée de E. Soit U la réunion de tous les idéaux de T. La réunion U est alors un idéal (la démonstration est analogue à celle donnée dans le cas de l'exemple 1). Toutefois, le point crucial est ici que U n'est pas égal à A. En effet, si $U = A$, alors $1 \in U$ et il y a par suite un idéal J de T tel que $1 \in J$ puisque U est la réunion des idéaux appartenant à T. Cela n'est pas possible puisque E est un ensemble d'idéaux *propres*. Par conséquent U est dans E, et est évidemment un majorant de T (et même une borne supérieure), et nous pouvons appliquer le lemme de Zorn pour conclure notre démonstration.

Soit V un espace vectoriel non nul sur un corps K. Soit $\{v_i\}_{i\in I}$ une famille d'éléments de V. Si $\{a_i\}_{i\in I}$ est une famille d'éléments de K, tels que $a_i = 0$ pour tous les indices; sauf pour un nombre fini d'entre eux, nous pouvons alors former la somme

$$\sum_{i\in I} a_i v_i.$$

Si $i_1, \ldots, i_n$ sont les indices pour lesquels $a_i \neq 0$, la somme ci-dessus est définie comme étant

$$a_{i_1} v_{i_1} + \cdots + a_{i_n} v_{\ n}.$$

Nous disons que cette famille $\{v_i\}_{i\in I}$ est *libre* si, chaque fois que nous avons une famille $\{a_i\}_{i\in I}$ où les $a_i \in K$ sont tous nuls sauf un nombre fini d'entre eux, et si

$$\sum_{i\in I} a_i v_i = 0,$$

alors tous les $a_i = 0$. Pour plus de simplicité, nous abrégeons «tous sauf un nombre fini d'entre eux» par «*presque tous*». On dit qu'une famille $\{v_i\}_{i\in I}$ d'éléments de V engendrent V si tout élément $v \in V$ peut s'écrire sous la forme

$$v = \sum_{i\in I} a_i v_i$$

pour une famille $\{a_i\}_{i\in I}$ d'éléments de K, presque tous les a_i étant nuls. Une famille $\{v_i\}_{i\in I}$ qui est libre et qui engendre V est dite base de V.

Si U est une partie de V, on peut considérer U comme une famille, indexée par ses propres éléments. Si l'on se donne un élément a_v pour chaque $v \in U$, et si ces éléments sont presque tous nuls, on peut former la somme

$$\sum_{v\in U} a_v v.$$

On peut de cette façon donner un sens au fait qu'une partie quelconque engendre V et est libre. On peut définir une *base* de V comme partie libre qui engendre V.

Théorème 2. *Soit V un espace vectoriel non nul sur le corps K. Il existe, dans ces conditions, une base de V.*

Démonstration. Soit E l'ensemble des parties libres de V. Cet ensemble E est non vide parce que pour tout $v \in V, v \neq 0$, l'ensemble $\{v\}$ est libre. Si B et B' sont des éléments de E, on pose $B \leqq B'$ si $B \leqq B'$. L'ensemble E est donc partiellement ordonné, et est inductif, puisque si T est une partie totalement ordonnée de E, alors

$$\bigcup_{B\in T} B$$

est un majorant de T dans E. Soit, d'après le lemme de Zorn, un élément maximal M de E. Soit $v \in V$. Puisque M est maximal, si $v \notin M$ l'ensemble $M \cup \{v\}$ n'est pas libre. Par suite, il existe des éléments $a_w \in K (w \in M)$ et $b \in K$, non tous nuls, tels que

$$bv + \sum_{w\in M} a_w w = 0.$$

Si $b = 0$, on contredit le fait que M est libre. Par suite, $b \neq 0$ et

$$v = \sum_{w\in M} - b^{-1} a_w w$$

est combinaison linéaire d'éléments de M. Si $v \in M$, alors v est trivialement combinaison linéaire d'éléments de M. Par suite M engendre V et est donc la base recherchée de V.

Exercices

1. Explicitez les démonstrations prouvant que les ensembles des exemples 2 et 3 sont inductifs.

2. Explicitez la démonstration de l'assertion faite au cours de la démonstration du théorème 2, et par laquelle il est affirmé que $\bigcup_{B\in T} B$ est une partie libre.

3. Soit A un anneau et E un module de type fini sur A, i.e. un module possédant un nombre fini de générateurs $v_1, \ldots, v_n$. Supposons que E n'est pas le module nul. Montrez que E possède un sous-module maximal, i.e. un sous-module $M \neq E$ tel que si N est un sous-module vérifiant $M \subset N \subset E$, alors $M = N$ ou $N = E$.

4. Soit A un anneau commutatif et E un sous-ensemble de A non vide. Montrez qu'il existe un idéal M dont l'intersection avec E est vide, et maximal pour cette propriété (on dit alors que M est un idéal maximal ne rencontrant pas E).

§3. *Nombres cardinaux*

Soient A et B des ensembles. On dira que le *cardinal de A est le même que le cardinal de B*, et on écrit

$$\operatorname{card}(A) = \operatorname{card}(B)$$

s'il existe une bijection de A sur B.

On dit que $\operatorname{card}(A) \leqq \operatorname{car}(B)$ s'il existe une application injective (injection) $f : A \to B$. On écrit aussi $\operatorname{card}(B) \geqq \operatorname{card}(A)$ dans ce cas. Il est clair que si $\operatorname{card}(A) = \operatorname{card}(B)$ et $\operatorname{card}(B) = \operatorname{card}(C)$, alors

$$\operatorname{card}(A) = \operatorname{card}(C).$$

Cela revient à dire que l'application composée de deux injections est injective. De la même façon, si $\operatorname{card}(A) = \operatorname{card}(B)$ et $\operatorname{card}(B) = \operatorname{card}(C)$, alors

$$\operatorname{card}(A) = \operatorname{card}(C).$$

Cela revient à dire que l'application composée de deux bijections est bijective. Enfin, on a de façon évidente $\operatorname{card}(A) = \operatorname{card}(A)$.

Nous allons d'abord étudier les ensembles dénombrables. Un ensemble D est dit *dénombrable* s'il existe une bijection entre D et les entiers positifs, bijection qui est appelée *énumération* de l'ensemble D.

Toute partie infinie d'un ensemble dénombrable est dénombrable.

On démontre facilement cela par récurrence. (Nous esquissons la démonstration: il suffit de montrer que toute partie infinie des entiers positifs est dénombrable. Soit $D = D_1$ une telle partie. Alors D_1 a au moins un élément a_1. Supposons que nous avons défini D_n par récurrence pour un entier $n \geqq 1$. Soit D_{n+1} l'ensemble de tous les éléments de D_{n+1} plus grands que D_n. Nous obtenons ainsi une application injective.

$$n \mapsto a_n$$

de $\mathbf{Z}^+$ dans D, dont on voit qu'elle est surjective.)

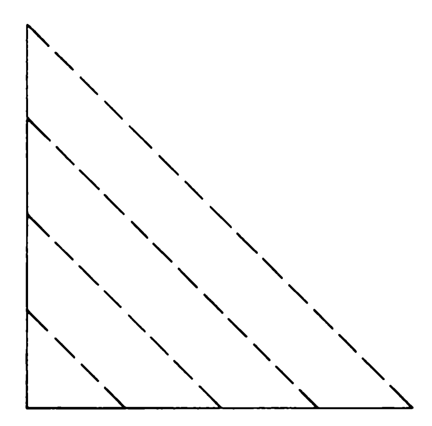
Théorème 3. *Si D est un ensemble dénombrable, $D \times D$ est également un ensemble dénombrable.*

Démonstration. Il suffit de démontrer que $\mathbf{Z}^+ \times \mathbf{Z}^+$ est dénombrable. Considérons l'application

$$(m, n) \mapsto 2^m 3^n.$$

C'est une application injective de $\mathbf{Z}^+ \times \mathbf{Z}^+$ dans $\mathbf{Z}^+$, et, par suite, $\mathbf{Z}^+ \times \mathbf{Z}^+$ a le même cardinal qu'une partie infinie de $\mathbf{Z}^+$, d'où la dénombrabilité de $\mathbf{Z}^+ \times \mathbf{Z}^+$, qu'on voulait montrer.

Nous avons utilisé dans cette démonstration la décomposition en nombres premiers des entiers. On peut également donner une démonstration du théorème 3 qui n'utilise pas ce fait. L'idée sous-tendant une telle démonstration est illustré par le schéma suivant:



Il nous faut trouver une bijection de $\mathbf{Z}^+$ dans $\mathbf{Z}^+ \times \mathbf{Z}^+$. Nous appliquons 1 sur (1, 1). Supposons par hypothèse de récurrence, que nous avons défini une application injective

$$f : \{1, .., n\} \to \mathbf{Z}^+ \times \mathbf{Z}^+.$$

On veut définir $f(n + 1)$.

Si $f(n) = (1, k)$, alors on pose $f(n + 1) = (k + 1, 1)$;

Si $f(n) = (r, k)$ avec $r \neq 1$, alors on pose $f(n + 1) = (r - 1, k + 1)$.

C'est alors une question de routine de vérifier qu'on obtient ainsi une injection de $\{1, ..., n + 1\}$ dans $\mathbf{Z}^+ \times \mathbf{Z}^+$. On obtient par récurrence, une application de $\mathbf{Z}^+$ dans $\mathbf{Z}^+ \times \mathbf{Z}^+$ qui est également bijective comme on le vérifie aisément. On peut décrire notre application f sur notre schéma de la façon suivante: partons du coin (1, 1), et déplaçons-nous à l'intérieur du quadrant, en partant de l'axe horizontal et en nous déplaçant en diagonale et vers la gauche jusqu'à ce que nous rencontrons l'axe vertical et de là, repartons d'un cran plus haut vers l'axe horizon, pour recommencer le processus. Il est géomètriquement clair que notre application passe par tous les points (i, j) de $\mathbf{Z}^+ \times \mathbf{Z}^+$.

***Corollaire* 1.** *Le produit $D \times \cdots \times D$ (n fois) est dénombrable, pour tout entier n positif.*

Démonstration. Par récurrence.

***Corollaire* 2.** *Soit* $\{D_1, D_2, ...,\}$ *une suite d'ensembles dénombrables, encore écrite* $\{D_i\}_{i\in}\mathbf{Z}^+$. *Dans ces conditions, la réunion*

$$U = \bigcup_{i=1}^{\infty} D_i$$

est dénombrable.

Démonstration. On a, pour chaque i, une énumération des éléments de D_i, par exemple

$$D_i = \{a_{i1}, a_{i2}, ...\}.$$

L'application

$$(i, j) \mapsto a_{ij}$$

est alors une application de $\mathbf{Z}^+ \times \mathbf{Z}^+$ dans **U**, et est, de fait, surjective. Soit

$$f : \mathbf{Z}^+ \times \mathbf{Z}^+ \to U$$

cette application. Il existe, pour tout $a \in U$, un élément $x \in \mathbf{Z}^+ \times \mathbf{Z}^+$ tel que $f(x) = a$. et on peut écrire cet élément x sous la forme x_a. La correspondance $a \mapsto x_a$ est une injection de U dans $\mathbf{Z}^+ \times \mathbf{Z}^+$, et on peut alors appliquer le théorème pour conclure.

Dans la démonstration précédente, nous avons utilisé un cas particulier d'un résultat sur les cardinaux qu'il est utile d'établir en général:

Soit $f : A \to B$ *une application surjective de* A *sur* B. *Alors*

$$\text{card}(B) \leqq \text{card}(A).$$

On le voit facilement, parce que, pour tout $y \in B$, il existe un élément $x \in A$, désigné par x_y, tel que $f(x_y) = y$. La correspondance $y \mapsto x_y$ est alors une application injective de B dans A, d'où par définition

$$\text{card}(B) \leqq \text{card(A)}.$$

Pour traiter de cardinaux quelconques, on a besoin d'un théorème peut-être moins trivial que dans le cas dénombrable.

***Théorème* 4** (Schroeder-Bernstein). *Soient* A *et* B *deux ensembles; supposons que* $\text{card}(A) \leqq \text{card}(B)$, *et que* $\text{card}(B) \leqq \text{card}(A)$. *Alors*

$$\text{card}(A) = \text{card}(B).$$

Démonstration. Soient $f : A \to B$ et $g : B \to A$ deux applications injectives. Divisons A en deux ensembles A_1 et A_2 disjoints. L'ensemble A_1 sera constitué de tous les $x \in A$ tels que, lorsque nous redescendons de x par une sucession d'applications réciproques

$$x, g^{-1}(x), f^{-1} \circ g^{-1}(x), g^{-1} \circ f^{-1} \circ g^{-1}(x), ...,$$

alors nous atteignons à une certaine étape un élément de A qui ne peut pas être redescendu par g dans B. Soit A_2 le complémentaire de A_1; en d'autres termes, soit A_2 l'ensemble des $x \in A$ qui peuvent être indéfiniment descendus, c'est-à-dire tels que nous soyons arrêtés dans B (i.e. que nous atteignons un élément de B qui n'a pas d'image réciproque par f dans A). On a alors $A = A_1 \cup A_2$. Nous allons définir une bijection h de A sur B.

Si $x \in A_1$, on pose $h(x) = f(x)$.

Si $x \in A_2$, on pose $h(x) = g^{-1}(x) =$ l'unique élément $y \in B$ tel que $g(y) = x$.

L'application h est alors injective de façon triviale. Il nous faut montrer que h est surjective. Soit $b \in B$. Si, lorsque nous essayons de redescendre b par une succession d'applications

$$\ldots \circ f^{-1} \circ g^{-1} \circ f^{-1} \circ g^{-1} \circ f^{-1}(b).$$

nous pouvons redescendre indéfiniment, ou si nous sommes arrêtés en B; alors $g(b)$ appartient à A_2 et, par suite, $b = h(g(b))$, de sorte que b se trouve dans l'image de h. D'autre part, si nous ne pouvons pas indéfiniment redescendre b, et restons arrêtés dans A, alors $f^{-1}(b)$ est défini (i.e. b est dans l'image de f), et $f^{-1}(b)$ se trouve dans A_1. Dans ce cas, $b = h(f^{-1}(b))$ est aussi dans l'image de h, comme il fallait le montrer.

Considérons maintenant des théorèmes concernant les sommes et les produits de cardinaux.

Nous allons réduire l'étude des cardinaux de produits d'ensembles quelconques, au produit d'ensembles dénombrables, en utilisant le lemme de Zorn. Remarquons d'abord qu'un ensemble infini contient toujours un ensemble dénombrable. En effet, puisque A est infini, nous pouvons commencer par choisir un élément $a_1 \in A$, et le complémentaire de $\{a_1\}$ est infini. Par hypothèse de récurrence, si nous avons choisi des éléments $a_1, \ldots, a_n$ dans A, le complémentaire de $\{a_1, \ldots, a_n\}$ est infini, et nous pouvons choisir a_{n+1} dans ce complémentaire. Nous obtenons de cette façon une suite d'éléments distincts de A, qui fournissent une partie dénombrable de A. (Pour une formalisation supplémentaire de cette démonstration, cf. Appendice, §3.)

Soit A un ensemble. On entend par *recouvrement* de A, un ensemble Γ de partie de A tel que la réunion

$$\bigcup_{C \in \Gamma} C$$

de tous les éléments de Γ soit égale à A. Nous disons que Γ est une *partition* si, lorsque C et C' de Γ sont distincts, ils sont disjoints.

Lemme. *Soit A un ensemble infini. Il existe une partition de A en ensembles dénombrables.*

Démonstration. Soit E l'ensemble dont les éléments sont des couples (B, Γ) formés d'une partie B de A, et d'une partition de B par des ensembles dénombrables. L'ensemble E n'est alors pas vide. En effet, puisque A est infini, A contient un

ensemble dénombrable D, et le couple $(D, \{D\})$ est dans E. Si (B, Γ) et (B', Γ) sont des éléments de E, on pose

$$(B, \Gamma) \leqq (B', \Gamma')$$

pour indiquer que $B \subset B'$, et que $\Gamma \subset \Gamma'$. Soit T une partie non vide totalement ordonnée de E. On peut écrire $T = \{(B_i, \Gamma_i)\}_{i \in I}$ pour un certain ensenble d'indices I. Soient

$$B = \bigcup_{i \in I} B_i \qquad \text{et} \qquad \Gamma = \bigcup_{i \in I} \Gamma_i.$$

Si $C, C' \in \Gamma$ et $C \neq C'$, il existe alors des indices i, j tels que $C \in \Gamma_i$ et $C' \in \Gamma_j$. Puisque T est totalement ordonné, nous avons, par exemple,

$$(B_i, \Gamma_i) \leqq (B_j, \Gamma_j).$$

Donc, en fait, C et C' sont tous deux éléments de Γ_j et, par conséquent, C et C' ont une intersection vide. D'autre part, si $x \in B$, alors $x \in B_i$ pour un i, et par suite, il existe $C \in \Gamma_i$ tel que $x \in C$. Par conséquent, Γ est une partition de B. Puisque les éléments de tout Γ_i sont des parties dénombrables de A, il s'ensuit que Γ est une partition de B par des ensembles dénombrables, qu'ainsi (B, Γ) est dans E et est un majorant de T. Par conséquent, E est inductif.

Soit (M, Δ) un élément maximal de E, dont l'existence est assurée par le lemme de Zorn. Supposons $M \neq A$. Si le complémentaire de M dans A est infini, il existe un ensemble dénombrable D contenu dans ce complémentaire. Alors

$$(M \cup D, \Delta \cup \{D\})$$

est un couple plus grand que (M, Δ) contredisant la maximalité de (M, Δ). Le complémentaire de M dans A est donc un ensemble fini F. Soit D_0 un élément de Δ. Soit $D_1 = \mathrm{D}_0 \cup F$. L'ensemble D_1 est donc dénombrable. Soit Δ_1 l'ensemble constitué de tous les éléments de Δ, excepté D_0, et de D_1. L'ensemble Δ_1 est un recouvrement de A par des ensembles dénombrables, comme il fallait le démontrer.

***Théorème* 5.** *Soit A un ensemble infini, et soit D un ensemble dénombrable; on a alors*

$$\operatorname{card}(A \times D) = \operatorname{card}(A).$$

Démonstration. On peut écrire, d'après le lemme,

$$A = \bigcup_{i \in I} D_i$$

comme partition d'ensembles dénombrables. Alors

$$A \times D = \bigcup_{i \in I} (D_i \times D).$$

Pour tout $i \in I$, il y a, d'après le théorème 3, une bijection de $D_i \times D$ sur D. Puisque les ensembles $D_i \times D$ sont disjoints, nous obtenons de cette façon une bijection de $A \times D$ sur A, comme souhaité.

***Corollaire* 1.** *Si F est un ensemble fini non vide, alors*

$$\text{card}(A \times F) = \text{card}(A).$$

Démonstration. Nous avons

$$\text{card}(A) \leqq \text{card}(A \times F) \leqq \text{card}(A \times D) = \text{card}(A).$$

Nous pouvons alors appliquer le théorème 4 pour obtenir le résultat voulu.

***Corollaire* 2.** *Soient A et B deus ensembles non vides, A étant fini; supposons que* $\text{card}(B) \leqq \text{card}(A)$. *On a alors*

$$\text{card}(A \cup B) = \text{card}(A).$$

Démonstration. On peut écrire $A \cup B = A \cup C$ pour une partie C de B telle que C et A sont disjoints. (On prend pour C l'ensemble de tous les éléments de B qui ne sont pas dans A.) On a alors $\text{card}(C) \leqq \text{car}(A)$. Nous pouvons donc construire une injection de $A \cup C$ dans le produit

$$A \times \{1, 2\}$$

de A avec un ensemble à deux éléments. Plus précisément, nous avons une bijection entre A et $A \times \{1\}$ d'une façon évidente, et aussi une injection de C dans $A \times \{2\}$. Par suite

$$\text{card}(A \cup C) \leqq \text{card}(A \times \{1, 2\}).$$

On achève la démonstration en utilisant le corollaire 1 et le théorème 4.

***Théorème* 6.** *Soit A un ensemble infini. On a alors*

$$\text{card}(A \times A) = \text{card}(A).$$

Démonstration. Soit E l'ensemble des couples (B, f) où B est une partie infinie de A et $f : B \to B \times B$ une bijection de B sur $B \times B$. L'ensemble E n'est pas vide puisque D est une partie dénombrable de A (dont nous connaissons l'existence) et qu'on peut toujours trouver une bijection de D sur $D \times D$. Si (B, f) et (B', f') sont dans E. (On pose $(B, f) \leqq (B', f')$ pour indiquer que $B \subset B'$ et que la restriction de f' à B est égale à f. L'ensemble E est alors partiellement ordonné, et nous affirmons que E est inductif. Soit T une partie non vide totalement ordonnée de E, et constituée, par exemple, des couples (B_i, f_i) pour des i appartenant à un certain ensemble I d'indices. Soit

$$M = \bigcup_{i \in I} B_i.$$

Nous allons définir une bijection $g : M \to M \times M$. Si $x \in M$, alors x appartient à un certain B_i. On pose $g(x) = f_i(x)$. Cette valeur $f_i(x)$ ne dépend pas du choix du B_i auquel appartient x. En effet, si $x \in B_j$, pour un certain $j \in I$, alors, par exemple,

$$(B_i, f_i) \leqq (B_j, f_j).$$

Par hypothèse, $B_i \subset B_j$, et $f_j(x) = f_i(x)$, de sorte que g est bien définie. Pour montrer que g est surjective, considérons $x, y \in M$ et $(x, y) \in M \times M$. Alors $x \in B_i$ pour un certain $i \in I$ et $y \in B_j$ pour un certain $j \in I$. Mais puisque T est totalement ordonnée, on a, par exemple, $(B_i, f_i) \leqq (B_j, f_j)$. Donc $B_i \subset B_j$, et $x, y \in B_j$. Il existe un élément $b \in B_j$ tel que $f_j(b) = (x, y) \in B_j \times B_j$. Par définition, $g(b) = (x, y)$, de sorte que g est surjective. Nous laissons le soin de démontrer que g est injective —achevant ainsi la démonstration du fait que g est une bijection— au lecteur. On voit alors que (M, g) est un majorant de T dans E, et donc que E est inductif.

Soit (M, g) un élément maximal de E, et soit C le complémentaire de M dans A. Si $\operatorname{card}(C) \leqq \operatorname{card}(A)$, alors

$$\operatorname{card}(M) \leqq \operatorname{card}(A) = \operatorname{card}(M \cup C) = \operatorname{card}(M),$$

d'après le corollaire 2 du théorème 5; par suite, $\operatorname{card}(M) = \operatorname{card}(A)$, d'après le théorème de Bernstein. Puisque $\operatorname{card}(M) = \operatorname{card}(M \times M)$, la démonstration est faite pour ce cas. Si $\operatorname{card}(M) \leqq \operatorname{card}(C)$, il existe une partie M_1 de C ayant même cardinal que M. Considérons

$$(M \cup M_1) \times (M \cup M_1) = (M \times M) \cup (M_1 \times M) \cup (M \times M_1) \cup (M_1 \times M_1).$$

D'après l'hypothèse faite sur M et le corollaire 2 du théorème 5, les trois derniers ensembles entre parenthèses du membre de droite de cette égalité ont même cardinal que M. Donc

$$(M \cup M_1) \times (M \cup M_1) = (M \times M) \cup M_2,$$

où M_2 est disjoint de $M \times M$ et de même cardinal que M. Définissons maintenant une bijection

$$g_1 : M \cup M_1 \to (M \cup M_1) \times (M \cup M_1).$$

Posons $g_1(x) = g(x)$ si $x \in M$, et prenons pour restriction de g_1 à M_1 n'importe quelle bijection de M_1 sur M_2. Nous avons de cette façon prolongé g à $M \cup M_1$, et le couple $(M \cup M_1, g_1)$ est dans E, contredisant ainsi la maximalité de (M, g). Le cas $\operatorname{card}(M) \leqq \operatorname{card}(C)$ ne peut donc pas avoir lieu, et notre théorème est démontré.

Corollaire 1. *Si A est un ensemble fini, et si $A^{(n)} = A \times \cdots \times A$ est le produit de n fois A, alors*

$$\operatorname{card}(A^{(n)}) = \operatorname{card}(A).$$

Démonstration. Par récurrence.

***Corollaire* 2.** *Si $A_1, \dots, A_n$ sont des ensembles non vides et si*

$$\text{card}(A_i) \leqq \text{card}(A_n)$$

pour $i = 1, \dots, n$, alors

$$\text{card}(A_1 \times \cdots \times A_n) = \text{card}(A_n).$$

Démonstration. On a

$$\text{card}(A_n) \leqq \text{card}(A_1 \times \cdots \times A_n) \leqq \text{card}(A_n \times \cdots \times A_n),$$

et on achève la démonstration, en utilisant le corollaire 2 et le théorème de Schroeder-Bernstein.

***Corollaire* 3.** *Soit A un ensemble infini, et soit Φ l'ensemble des parties finies de A. Alors*

$$\text{card}(\Phi) = \text{card}(A).$$

Démonstration. Soit Φ_n l'ensemble des parties de A ayant exactement n éléments, pour tout entier $n = 1, 2, \dots$ Montrons d'abord que card $(\Phi_n) \leqq$ card (A). Si F ets un élément de Φ_n, on ordonne les éléments de F de n'importe quelle façon; on pose, par exemple,

$$F = \{x_1, \dots, x_n\},$$

et on associe l'élément $(x_1, \dots, x_n) \in A^{(n)}$ à F, par

$$F \mapsto (x_1, \dots, x_n).$$

Si G est une autre partie de A à n éléments, par exemple si $G = \{y_1, \dots, y_n\}$, et si $G \neq F$, alors

$$(x_1, \dots, x_n) \neq (y_1, \dots, y_n).$$

Par conséquent, notre application

$$F \mapsto (x_1, \dots, x_n)$$

de Φ_n dans $A^{(n)}$ est injective. On en conclut, d'après le corollaire 1, que

$$\text{card}(\Phi_n) \leqq \text{card}(A).$$

Mais les Φ_n, pour $n = 1, 2, \dots$, forment une partition de Φ, et on montre en exercice que card$(\Phi) \leqq$ card(A) (cf. exercice 1). Puisque

$$\text{card}(A) \leqq \text{card}(\Phi),$$

(car, en particulier, card(Φ_1) = card(A)), on voit que notre corollaire est démontré.

Nous allons voir, au théorème suivant, qu'étant donné un ensemble, il en existe toujours un autre de cardinal supérieur.

Théorème 7. *Soit A un ensemble infini et soit T l'ensemble à deux éléments* $\{0, 1\}$. *Soit F l'ensemble des applications de A dans T. Alors*

$$\text{card(A)} \leqq \text{card}(M)\,\text{et}\,\text{card}(A) \neq \text{card}(M).$$

Démonstration. On définit, pour tout $x \in A$,

$$f_x : A \to \{0, 1\}$$

comme étant l'application telle que $f_x(x) = 1$ et $f_x(y) = 0$ si $y \neq x$. L'application $x \mapsto f_x$ est évidemment une injection de A dans M, de telle sorte que $\text{card}(A) \leqq$ $\leqq \text{card}(M)$. Supposons que $\text{card}(A) = \text{card}(M)$. Soit

$$x \mapsto g_x$$

une bijection entre A et M. On définit une application $h : A \to \{0, 1\}$ en posant:

$$h(x) = 0, \quad \text{si} \quad g_x(x) = 1,$$
$$h(x) = 1, \quad \text{si} \quad g_x(x) = 0.$$

On a certainement $h \neq g_x$ pour tout x, ce qui contredit l'hypothèse que $x \mapsto g_x$ est une bijection et prouve le théorème 7.

Corollaire *Soit A un ensemble infini, et E l'ensemble des parties de A. Alors*

$$\text{card}(A) \leqq \text{card}(E) \qquad \text{et} \qquad \text{card}(A) \neq \text{card}(E).$$

Démonstration. Nous la laissons en exercice. [*Indication:* si B est une partie non vide, utilisez la fonction caractéristique φ_B telle que

$$\varphi_B(x) = 1, \quad \text{si} \quad x \in B,$$
$$\varphi_B(x) = 0, \quad \text{si} \quad x \notin B.$$

Que pouvez-vous dire de la correspondance $B \mapsto \varphi_B$?]

Exercices

1. Démontrez l'assertion faite dans la démonstration du corollaire 3 du théorème 6.
2. Soient A un ensemble infini, et Φ_n l'ensemble des parties de A à n éléments. Montrez que

$$\text{card}\,(A) \leqq \text{card}\,(\Phi_n)$$

pour $n \geqq 1$.
3. Soient $A_1, A_2, \ldots$ des ensembles infinis; supposons que

$$\text{card}\,(A_i) \leqq \text{card}\,(A)$$

pour un ensemble A et pour tout i. Montrez que

$$\text{card}\,\Big(\bigcup_{i=1} A_i\Big) \leqq \text{card}\,(A).$$

4. Soit K un sous-corps du corps des nombres complexes. Montrez que pour tout entier $n \geqq 1$, le cardinal de l'ensemble des extensions de K de degré n dans $\mathbf{C}$ est $\leqq \text{card}\,(K)$.

5. Soient K un corps infini, et E une extension algébrique de K. Montrez que card (E) = card (K).

6. Terminez la démonstration du corollaire du théorème 7.

7. Si A et B sont des ensembles, on désigne par $F(A, B)$ l'ensemble de toutes les applications de A dans B. Si B et B' sont de même cardinal, montrez que $F(A, B)$ et $(F(A, B')$ ont même cardinal. Montrez que $F(A, B)$ et $F(A', B)$ ont même cardinal si A et A' ont même cardinal.

8. Soit A un ensemble infini et notons en abrégé α, le cardinal de A. On abrège card(B) en β pour un ensemble infini B. On pose $\alpha\beta = \text{card}(A \times B)$. Soit B' un ensemble disjoint de A tel que card(B) = card(B'). On pose $\alpha + \beta = \text{card}(A \cup B')$. On désigne par B^A l'ensemble de toutes les applications de A dans B, et on désigne card(B^A) par β^α. Soit γ l'abréviation de card(C) pour un ensemble C infini. Démontrez les assertions suivantes:

(a) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$; (b) $\alpha\beta = \beta\alpha$; (c) $\alpha^{\beta + \gamma} = \alpha^\beta\alpha^\gamma$.

9. Soit K un corps infini. Démontrez qu'il existe un corps algébriquement clos $\overline{K}$ contenant K comme sous-corps et algébrique sur K. [*Indication:* soit Ω un ensemble de cardinal strictement supérieur au cardinal de K, et contenant K. Considérez l'ensemble C de tous les couples (E, φ), où E est une partie de Ω telle que $K \subset E$, et où φ désigne une loi d'addition et de multiplication sur E qui en fait un corps dont K est un sous-corps, et telle que E soit algébrique sur K. On définit un ordre partiel sur C d'une façon évidente; montrez que C est inductif, et qu'un élément maximal de C est algébrique sur K et algébriquement clos. Vous aurez besoin de l'exercice 5 à la dernière étape de la démonstration.]

10. Soit K un corps infini. Montrez que le corps des fractions rationnelles $K(t)$ est de même cardinal que K.

11. Soit J_n l'ensemble des entiers $\{1, ..., n\}$. Soit $\mathbf{Z}^+$ l'ensemble des entiers positifs. Montrez que les ensembles suivants ont même cardinal:

(a) l'ensemble $F(\mathbf{Z}^+, J_n)$ de toutes les applications de $\mathbf{Z}^+$ dans J_n,
(b) l'ensemble $F(\mathbf{Z}^+, J_2)$ de toutes les applications de $\mathbf{Z}^+$ dans J_2,
(c) l'ensemble de tous les nombres réels x tels que $0 \leqq x < 1$,
(d) l'ensemble de tous les réels.
[*Indication:* utilisez les développements décimaux.]

12. Si vous connaissez la représentation des réels par les fractions continues, montrez que $F(\mathbf{Z}^+, \mathbf{Z}^+)$ a même cardinal que l'ensemble des nombres réels.

§4. *Ensembles bien ordonnés*

On dit qu'un ensemble A est *bien ordonné*, s'il est totalement ordonné, et si toute partie B non vide de A possède un plus petit élément, i.e. un élément $a \in B$ tel que $a \leqq x$ pour tout $x \in B$.

Exemple 1. L'ensemble $\mathbf{Z}^+$ des entiers positifs est bien ordonné. Tout ensemble fini peut être bien ordonné, et un ensemble dénombrable peut être bien ordonné: une bijection quelconque de D avec $\mathbf{Z}^+$ fera de D un ensemble bien ordonné.

Exemple 2. Soit D un ensemble dénombrable bien ordonné. Soit b un élément d'un certain ensemble, tel que $b \notin D$. Soit $A = D \cup \{b\}$. On pose $x \leqq b$, pour tout $x \in D$. L'ensemble A est alors totalement ordonné et, en fait, il est même bien ordonné.

Démonstration. Soit B une partie non vide de A. Si B est réduite à $\{b\}$, alors b est le plus petit élément de B. Sinon B contient un élément $a \in D$. L'intersection $B \cap D$ n'est alors pas vide, et par suite possède un plus petit élément, qui est évidemment un plus petit élément de B.

Exemple 3. Soient D_1 et D_2 deux ensembles dénombrables, tous deux bien ordonnés; supposons que $D_1 \cap D_2$ est vide. Soit $A = D_1 \cup D_2$. On définit un ordre total dans A en posant $x < y$ pour tout $x \in D_1$ et tout $y \in D_2$. En utilisant le même genre de raisonnement qu'à l'exemple 2, on voit que A est bien ordonné.

Exemple 4. Etant donné une suite d'ensembles dénombrables disjoints, on définit $A = \cup D_i$ par récurrence, et on définit un ordre qui fait de A un ensemble bien ordonné en ordonnant chaque D_i comme $\mathbf{Z}^+$, et en posant $x < y$ lorsque $x \in D_i$ et $y \in D_{i+1}$. On peut illustrer cette situation comme suit

|———————	|———————	|———————
D_1	D_2	D_3

***Théorème* 8.** *Tout ensemble infini non vide peut être bien ordonné.*

Démonstration. Soit E l'ensemble de tous les couples (X, R) où X est une partie de A et R un ordre total qui fasse de X un ensemble bien ordonné. L'ensemble E n'est pas vide, lorsque, étant donnée une partie D dénombrable de A, on peut toujours en faire une partie bien ordonnée en l'ordonnant comme les entiers positifs. Si (X, R) et (Y, Q) sont éléments de E, on pose $(X, R) \leqq (Y, Q)$ si $X \subset Y$, si la rectriction de Q à X est égale à R et si tout élément $y \in Y$, $y \notin X$ est tel que $x < y$ pour tout $x \in X$. L'ensemble E est alors partiellement ordonné. Pour montrer que E est inductif, considérons une partie T non vide et totalement ordonnée de E, et posons, pour fixer les idées, $T = \{(X_i, R_i)\}_{i \in I}$. Soit

$$M = \bigcup_{i \in I} X_i.$$

Soient $x, y \in M$. Il existe $i, j \in I$ tels que $x \in X_i$ et $y \in X_j$. Puisque T est totalement ordonné, on a, par exemple, $(X_i, R_i) \leqq (X_j, R_j)$. Les deux éléments x et y sont donc dans X_j. On pose $x \leqq y$ dans M si $x \leqq y$ dans X_j. On voit facilement que cette définition est indépendante du choix du (X_j, R_j) dans lequel sont x et y, et on vérifie de façon triviale qu'on a ainsi défini un ordre total sur M, que nous désignons par (M, P). Nous affirmons que M est bien ordonné par cet ordre total. Considérons, pour le voir, une partie N non vide de M. Soit $x_0 \in N$. Il existe alors un certain $i_0 \in I$ tel que $x_0 \in X_{i_0}$. La partie $N \cap X_{i_0}$ n'est pas vide. Soit a un minorant de cette partie. Nous affirmons que a est en fait le plus petit élément de N. Soit $x \in N$. L'élément x appartient donc à un certain X_i. Puisque T est totalement ordonné, on a

$$(X_i, R_i) \leqq (X_{i_0}, R_{i_0}) \quad \text{ou} \quad (X_{i_0}, R_{i_0}) \geqq (X_i, R_i).$$

Dans le premier cas, $x \in X_i \subset X_{i_0}$ et, par suite, $a \leqq x$. Dans le second, si $x \notin X_{i_0}$, alors $a \in x$ par définition. Cela démontre que (M, P) est bien ordonné.

Nous avons par conséquent démontré que E est inductif. Il existe, d'après le lemme de Zorn, un élément maximal (M, P) dans E. L'ensemble M est bien ordonné, et il ne reste plus qu'à montrer que $M = A$. Supposons $M \neq A$, et soit z un élément de A tel que $z \notin M$. Soit $M' = M \cup \{z\}$. On définit un ordre total sur M' en posant $x < z$, pour tout $x \in M$. L'ensemble M' est alors bien ordonné: considérons, pour le voir, une partie N non vide totalement ordonnée de M'. Si $N \cap M$ n'est pas vide, alors $N \cap M$ a un plus petit élément, qui est évidemment un plus petit élément de N. Si $N \cap M$ est vide, alors $N = \{z\}$, et alors z lui-même est un plus petit élément de N. Cela contredit le fait que M est maximal dans E. D'où $M = A$; notre théorème est démontré.

§5. *Démonstration du lemme de Zorn*

Le lemme de Zorn n'est pas intellectuellement tout à fait satisfaisant comme axiome, parce que ce qu'il affirme est trop compliqué, et parce qu'on ne peut pas aisément concrétiser l'existence de l'élément maximal assurée par son énoncé. Nous allons montrer dans ce paragraphe comment on peut déduire le lemme de Zorn d'autres propriétés des ensembles, que chacun trouve tout de suite intellectuellement acceptables.

A partir de maintenant, et jusqu'à la fin de la démonstration du théorème 9, A est un ensemble non vide partiellement ordonné et strictement inductif. Rappelons que *strictement inductif* signifie que toute partie non vide totalement ordonnée possède une borne supérieure. Supposons que nous nous soyons donnés une application $f : A \to A$ telle que pour tout $x \in A$, nous avons $x \leqq f(x)$.

Soit $a \in A$. Soit B une partie de A. On dit que B est *admissible* si

(1) B contient a,

(2) on a $f(B) \subset B$,

(3) la borne supérieure de toute partie T totalement ordonnée de A est dans B.

La partie B est aussi strictement inductive pour l'ordre induit par A. Nous allons le démontrer:

Théorème 9 (Bourbaki). *Soit A un ensemble non vide partiellement ordonné et strictement inductif. Soit $f : A \to A$ une application telle que, pour tout $x \in A$, $x \leqq f(x)$. Il existe alors un élément $x_0 \in A$ tel que $f(x_0) = x_0$.*

Démonstration. Supposons que A soit totalement ordonné. Il y a alors, par hypothèse, une borne supérieure $b \in A$, et on a alors

$$b \leqq f(b) \leqq b,$$

de sorte que dans ce cas, notre théorème est clairement prouvé. Tout le problème va être de ramener le théorème à cette situation. Autrement dit, nous avons besoin de trouver une partie admissible totalement ordonnée de A.

Si nous excluons de A tous les éléments $x \in A$ tels que x ne soit pas $\geqq a$, les

éléments restants constituent évidemment une partie admissible. On peut alors supposer, sans restreindre la généralité des hypothèses, que A possède un plus petit élément a, c'est-à-dire que $a \leqq x$ pour tout $x \in A$.

Soit M l'intersection de toutes les parties admissibles de A. Remarquons que A lui-même est une partie admissible, et que toutes les parties admissibles de A contiennent a, de sorte que M n'est pas vide. De plus, M est lui-même une partie admissible de A. Pour voir cela, considérons $x \in M$. L'élément x est alors dans toute partie admissible, si bien que $f(x)$ est aussi dans toute partie admissible et que, par suite, $f(x) \in M$. Par conséquent $f(M) \subset M$. Si T est une partie non vide totalement ordonnée de M, et si b est la borne supérieure de T dans A, alors b se trouve dans toute partie admissible de A, et donc se trouve dans M. Il s'ensuit que M est la plus petite partie admissible de A, telle que toute partie admissible de A contenue dans M soit égale à M.

Nous allons démontrer que M est totalement ordonné, et prouver par là le théorème 9.

[Nous faisons tout d'abord quelques remarques qui n'appartiennent pas à la démonstration, mais aident à comprendre les lemmes qui suivent. Puisque $a \in M$, on voit que $f(a) \in M$, que $f \circ f(a) \in M$, et qu'en général $f^n(a) \in M$. De plus

$$a \leqq f(a) \leqq f^2(a) \leqq \dots$$

Si nous avons une égalité quelque part dans cette suite d'inégalités, nous en avons terminé, aussi pouvons-nous supposer que les inégalités sont toutes strictes. Soit D_0 l'ensemble totalement ordonné $\{f^n(a)\}_{n \geqq 0}$. L'ensemble D_0 a donc l'allure suivante:

$$a < f(a) < f^2(a) < \cdots < f^n(a) < \cdots$$

Soit a_1 la borne supérieure de D_0. On peut alors former

$$a_1 < f(a_1) < f^2(a_1) \dots$$

de façon à obtenir D_1, et on peut poursuivre ce procédé pour obtenir

$$D_1, D_2, \dots$$

Il est clair que $D_1, D_2, \dots$ sont contenus dans M. Si nous savions exprimer de façon précise le fait qu'on peut construire une chaîne sans fin de tels ensembles dénombrables, nous pourrions obtenir ce que nous cherchons. Le noeud de la question est que nous sommes actuellement en train de démontrer le lemme de Zorn, lequel est l'outil le plus naturel à utiliser pour assurer l'existence d'une telle chaîne. Remarquons cependant qu'étant donnée une telle chaîne, ses éléments possèdent deux propriétés: si c est un élément d'une telle chaîne et si $x \leqq c$, alors $f(x) \leqq c$. De plus, il n'y a aucun élément entre c et $f(x)$, et si x est un élément de la chaîne, alors $x \leqq c$ ou $f(c) \leqq x$. Nous allons démontrer deux lemmes qui montrent que les éléments de M possèdent ces propriétés.]

Soit $c \in M$. Nous disons que c est un *extremum* de M si lorsque $x \in M$ et $x < c$, alors $f(x) \leqq c$. Pour tout extremum $c \in M$, posons

$$M_c = \{x \in M \quad \text{tels que} \quad x \leqq c \quad \text{ou} \quad f(c) \leqq x\}.$$

Remarquons que M_c n'est pas vide car a lui appartient.

Lemme 1. *On a $M_c = M$, pour tout extremum c de M.*

Démonstration. Il va suffire de démontrer que M_c est une partie admissible. Soit $x \in M_c$. Si $x < c$ alors $f(x) \leqq c$, de sorte que $f(x) \in M_c$. Si $x = c$, alors $f(x) = f(c)$ dans M_c. Si $f(c) \leqq x$, alors $f(c) \leqq x \leqq f(x)$, de telle sorte qu'une fois de plus $f(x) \in M_c$. Nous avons donc démontré que $f(M_c) \subset M_c$.

Soit T une partie totalement ordonnée de M_c, et soit b la borne supérieure de T dans M. Si tous les éléments $x \in T$ sont $\leqq c$, alors $b \leqq c$ et $b \in M_c$. S'il existe $x \in T$ tel que $f(c) \leqq x$, alors $f(c) \leqq x \leqq b$, de sorte que b est dans M_c. Cela démontre notre lemme.

Lemme 2. *Tout élément de M est un extremum.*

Démonstration. Soit E l'ensemble des extremums (ou extrema) de M. L'ensemble E n'est alors pas vide, car $a \in E$. Il va suffire de démontrer que E est une partie admissible. Démontrons d'abord que $f(E) \subset E$. Soit $c \in E$. Soit $x \in M$ et supposons que $x < f(c)$. D'après le lemme 1, $M = M_c$ et par suite, on a soit $x < c$, soit $x = c$, soit $f(c) \leqq x$. Cette dernière éventualité ne peut pas avoir lieu, parce que $x < f(c)$. Si $x < c$, alors $f(x) \leqq c \leqq f(c)$.

Si $x = c$, alors $f(x) = f(c)$, d'où $f(E) \subset E$. Enfin, la partie T de E est totalement ordonnée. Soit b sa borne supérieure dans M. Soit $x \in M, x < b$. On doit montrer que $f(x) \leqq b$. On sait, d'après le lemme 1, que pour tout $t \in T$, on a $M_t = M$ d'où $x \leqq t$ ou $f(t) \leqq x$. Si, pour un $t \in T$, on a $x \leqq t$, alors $f(x) \leqq f(t) \leqq b$ et nous avons terminé. Sinon on a $f(t) \leqq x$ pour tout $t \in T$, d'où le fait que c est un majorant de T et l'inégalité impossible $b \leqq x$. Cela démontre que E est admissible, et notre lemme est démontré.

Nous voyons maintenant de façon triviale que M est totalement ordonnée. Soient en effet $x, y \in M$. Le point x est un extremum de M d'après le lemme 2, et $y \in M_x$, si bien que $y \leqq x$ ou

$$x \leqq f(x) \leqq y,$$

démontrant ainsi que M est totalement ordonnée. Comme on l'a remarqué précédemment cela achève la démonstration du théorème 9.

Nous allons obtenir le lemme de Zorn essentiellement comme corollaire du théorème 9 et nous l'obtenons tout d'abord sous la forme légèrement plus faible qui suit:

Corollaire 1. *Soit A un ensemble non vide et strictement inductif. Il possède un élément maximal.*

Démonstration. Supposons que A n'a pas d'élément maximal. Il existe alors pour tout $x \in A$ un élément $y_x \in A$ tel que $x < y_x$. Soit $f : A \to A$ l'application telle que $f(x) = y_x$ pour tout $x \in A$. L'ensemble A et f vérifient alors les hypothèses du théorème 9, et l'application du théorème 9 mène à une contradiction.

La seule différence entre le corollaire 1 et le lemme de Zorn réside en ce que, dans le corollaire 1, on suppose qu'une partie non vide totalement ordonnée possède une borne supérieure (un plus petit majorant, N.d.T.), au lieu d'un majorant. C'est, cependant, fort simple de ramener le lemme de Zorn à la forme d'apparence plus faible du corollaire 1. C'est ce que nous faisons au corollaire 2.

***Corollaire* 2** (lemme de Zorn). *Soit E un ensemble non vide et inductif. Il possède un élément maximal.*

Démonstration. Soit A l'ensemble de toutes les parties non vides totalement ordonnées de E. L'ensemble A n'est pas vide, puisque toute partie de E réduite à un élément appartient à A. Si $X, Y \in A$, on pose $X \leqq Y$ pour indiquer que $X \subset Y$. L'ensemble A est alors partiellement ordonné, et est en fait strictement inductif. Soit en effet $T = \{X_i\}_{i\in I}$ une partie totalement ordonnée de A. Soit

$$Z = \bigcup_{i\in I} X_i.$$

L'ensemble Z est alors totalement ordonné. Considérons, pour le voir, les éléments x et y de Z. On a alors $x \in X_i$ et $y \in X_j$ pour des indices i et j de I. Puisque T est totalement ordonnée, on a, par exemple $X_i \subset X_j$. Par suite $x, y \in X_j$, et, puisque X_j est totalement ordonné, $x \leqq y$ ou $y \leqq x$. L'ensemble Z est alors totalement ordonné, et est évidemment une borne supérieure de T dans A. On tire alors du corollaire 2 le fait que A possède un élément X_0 maximal. Cela veut dire que X_0 est une partie totalement ordonnée maximale de E (non vide). Soit m la borne supérieure de X_0 dans E. L'élément m est l'élément maximal cherché de E. En effet, si $x \in E$ et si $m \leqq x$, alors $X_0 \cup \{x\}$ est totalement ordonnée, donc égale à X_0 du fait de la maximalité de X_0. On a donc $x \in X_0$ et $x \leqq m$. Par suite $x = m$, comme on veut le démontrer.

Remarque. Le niveau auquel nous sommes arrivés dans les raisonnements de ce chapitre permet une axiomatisation déjà élaborée de la théorie des ensembles. Puisque tous les raisonnements faits dans ce chapitre sont aisément acceptables pour des mathématiciens professionnels, il est de bonne politique de nous en tenir là, sans plus essayer de regarder les choses de façon plus fondamentale.

On peut, cependant, s'intéresser à ces fondements en eux-mêmes si on en a le goût. Nous renvoyons le lecteur amateur du sujet à des livres techniques traitant de ces questions. Nous faisons quand même une remarque supplémentaire sur la démonstration du lemme de Zorn. Lorsque le lecteur relit le corollaire 1, il constate que nous disons: «soit $f : A \to A$ une application telle que $f(x) = y_x$». Lorsque l'on pose les fondements de la théorie des ensembles, on a besoin d'un axiome particulier pour assurer l'existence d'une telle application. Cet axiome s'appelle l'*axiome du choix*, si bien que ce que nous avons démontré, c'est que l'axiome du choix implique le lemme de Zorn. Indépendamment énoncé, l'axiome du choix dit ceci:

Soit $\{E_i\}_{i\in I}$ une famille d'ensembles, tous supposés non vides. Il existe alors une famille d'éléments $\{x_i\}_{i\in I}$, où chaque $x_i \in E_i$.

Nous avons défini, au corollaire 2, et pour tout $x \in A$, la partie B_x comme étant l'ensemble des $y \in A$ tel que $x < y$. Si aucun B_x n'est vide et en prenant $A = I$, on obtient l'existence de la famille $\{y_x\}_{x \in A}$ par l'axiome du choix. Mais, comme on l'a dit au début du paragraphe, pour qui fait-elle problème?

Appendice

§1. *Les entiers naturels*

Le but de cet appendice est de montrer comment on peut obtenir axiomatiquement les entiers, en utilisant exclusivement le vocabulaire et les propriétés élémentaires des ensembles.

Supposons qu'on se soit donné une fois pour toutes un ensemble $\mathbf{N}$ appelé *ensemble des entiers naturels*, et une application $\sigma : \mathbf{N} \rightarrow \mathbf{N}$, vérifiant les axiomes suivants (de Peano):

EN 1. *Il y a un élément* $0 \in \mathbf{N}$.
EN 2. *On a* $\sigma(0) \neq 0$, *et si on désigne par* $\mathbf{N}^+$ *l'ensemble des* $n \in \mathbf{N}$ *tels que* $n \neq 0$, *alors l'application* $x \mapsto \sigma(x)$ *est une bijection entre* $\mathbf{N}$ *et* $\mathbf{N}^+$.
EN 3. *Si* P *est une partie de* $\mathbf{N}$, *si* $0 \in P$ *et si* $\sigma(n) \in P$ *lorsque* $n \in P$, *alors* $P = \mathbf{N}$.

On désigne souvent $\sigma(n)$ par n' et on imagine n' comme le successeur de n. Le lecteur a reconnu dans EN 3 l'axiome de récurrence.

Désignons $\sigma(0)$ par 1.

Notre prochain travail va être de définir une addition et une multiplication entre entiers naturels.

Lemme 1. *Soient* $f : \mathbf{N} \rightarrow \mathbf{N}$ *et* $g : \mathbf{N} \rightarrow \mathbf{N}$ *deux applications telles que*

$$f(0) = g(0) \text{ et } \begin{cases} f(n') = f(n)' \\ g(n') = g(n)'\,; \end{cases}$$

dans ces conditions, $f = g$.

Démonstration. Soit P la partie de $\mathbf{N}$ formée des n tels que $f(n) = g(n)$. Comme P satisfait évidemment les conditions de EN 3, on a $P = \mathbf{N}$, ce qui démontre le lemme.

Pour tout $m \in \mathbf{N}$, on veut définir $m + n$ pour $n \in \mathbf{N}$, de telle sorte que

$$(1_m) \qquad m + 0 = m \qquad \text{et} \qquad m + n' = (m + n)', \qquad \text{pout tout } n \in \mathbf{N}.$$

D'après le lemme, cela n'est possible que d'une façon.

Si $m = 0$, on pose $0 + n = n$ pour tout $n \in \mathbf{N}$. L'égalité 1_m est donc évidemment vérifiée. Soit Q l'ensemble des $m \in \mathbf{N}$ pour lesquels on peut définir $m + n$ pour tout $n \in \mathbf{N}$ de façon à ce que (1_m) soit vérifiée. Alors $0 \in Q$. Supposons que $m \in Q$. On pose pour tout $n \in \mathbf{N}$.

$$m' + 0 = m' \qquad \text{et} \qquad m' + n = (m + n)'.$$

On a alors

$$m' + n' = (m + n')' = \big((m + n)'\big)' = (m' + n)'.$$

Par suite $(1_{m'})$ est vérifiée et $m' \in Q$. Cela démontre que $Q = \mathbf{N}$, et que nous avons bien défini l'addition pour tous les couples (m, n) d'entiers naturels.

On démontre aisément les propriétés de l'addition.

Commutativité. Soit E l'ensemble des entiers naturels m tels que

$$(2_m) \qquad m + n = n + m \qquad \text{pour tout} \quad n \in \mathbf{N}.$$

L'élément 0 est évidemment dans E, et si $m \in E$, alors

$$m' + n = (m + n)' = (n + m)' = n + m',$$

ce qui démontre que $E = \mathbf{N}$, comme on le voulait.

Associativité. Soit E l'ensemble des entiers naturels tels que

$$(3_m) \qquad (m + n) + k = m + (n + k), \qquad \text{pour tous } n, k \in \mathbf{N}.$$

L'élément 0 est alors dans E. Supposons que $m \in E$. On a alors

$$\begin{aligned}(m' + n) + k &= (m + n)' + k = \big((m + n) + k\big)' \\ &= \big(m + (n + k)\big)' = m' + (n + k),\end{aligned}$$

ce qui démontre que $E = \mathbf{N}$, comme on le voulait.

Régularité. Soit m un entier naturel. Nous disons que l'élément m est *régulier*, si, pour tous $k, n \in \mathbf{N}$ vérifiant $m + k = m + n$, on a $k = n$. Soit E l'ensemble des éléments réguliers. L'élément $0 \in E$, et si $m \in E$, alors

$$m' + k = m' + n \quad \text{implique} \quad (m + k)' = (m + n)'.$$

Puisque l'application $x \mapsto x'$ est injective, il s'ensuit que $m + k = m + n$, d'où $k = n$. En appliquant EN 3, on trouve $E = \mathbf{N}$.

Pour l'étude de la multiplication et d'autres propriétés, il nous faut généraliser le lemme 1.

Lemme 2. *Soit E un ensemble, et soit $\varphi : E \to E$ une application de E dans lui-même. Soient f et g deux applications de $\mathbf{N}$ dans E. Si*

$$f(0) = g(0) \quad \text{et} \quad \begin{cases} f(n') = \varphi \circ f(n), \\ g(n') = \varphi \circ g(n), \end{cases}$$

pour tout $n \in \mathbf{N}$, alors $f = g$.

Démonstration. Triviale par récurrence.

Il résulte du lemme 2 qu'il y a, pour tout entier naturel m, au plus une façon de définir un produit mn satisfaisant à

$$m0 = 0 \qquad \text{et} \qquad mn' = mn + m,$$

pour tout $n \in \mathbf{N}$. Nous définissons en fait le produit de la même façon que nous l'avons fait pour l'addition, c'est-à-dire par récurrence; on démontre alors d'une façon analogue que ce produit est commutatif, associatif et distributif, i.e. tel que

$$m(n + k) = mn + mk,$$

pour tous $m, n, k \in \mathbf{N}$. Nous laissons les détails au lecteur.

On obtient ainsi toutes les propriétés d'anneau, *à l'exception de* celle de groupe additif: il nous manque les inverses additifs (qui sont les opposés, N.d.T.). Remarquez que 1 est un élément neutre pour la multiplication, c'est-à-dire que $1m = m$ pour tout $m \in \mathbf{N}$.

Il n'est pas aussi facile de démontrer que les éléments non nuls sont réguliers pour la multiplication, c'est-à-dire que si $mk = mn$ et $m \neq 0$, alors $k = n$. Nous laissons encore cela au lecteur. Remarquons, en particulier, que si $mn \neq 0$, alors $m \neq 0$ et $n \neq 0$.

Rappelons qu'un *ordre* sur un ensemble X est une relation $x \leqq y$ entre certains couples (x, y) d'éléments de X, satisfaisant les conditions (pour tous $x, y, z \in X$):

OP 1. *On a* $x \leqq x$.
OP 2. *Si* $x \leqq y$ *et* $y \leqq z$, *alors* $x \leqq z$.
OP 3. *Si* $x \leqq y$ *et* $y \leqq x$, *alors* $x = y$.

L'*ordre* est dit *total* si, étant donné $x, y \in X$, nous avons $x \leqq y$ ou $y \leqq x$. On écrit $x < y$ si $x \leqq y$ et $x \neq y$.

On peut définir un ordre sur $\mathbf{N}$ en posant $n \leqq m$ s'il existe $k \in \mathbf{N}$ tel que $m = n + k$. La démonstration du fait qu'il s'agit d'un ordre est de routine et laissée au lecteur. *C'est là en fait un ordre total,* et nous allons le prouver. Etant donné un nombre naturel m, soit C_m l'ensemble des $n \in \mathbf{N}$ tel que $n \leqq m$ ou $m \leqq n$. L'élément 0 appartient certainement à C_m. Supposons que $n \in C_m$. Si $n = m$, alors $n' = m + 1$, de sorte que $m \leqq n'$. Si $n < m$, alors $m = n + k'$ pour un $k \in \mathbf{N}$, de telle sorte que

$$m = n + k' = (n + k)' = n' + k,$$

et $n' \leqq m$. Si $m \leqq n$, alors pour un certain k, on a $n = m + k$, de sorte que $n + 1 = m + k + 1$ et $m \leqq n + 1$. D'après EN 3, $C_m = \mathbf{N}$, prouvant par là que notre ordre est total.

Il est maintenant facile de démontrer les propositions habituelles concernant les inégalités comme

$$m < n \quad \text{si et seulement si} \quad m + k < n + k \quad \text{pour un} \quad k \in \mathbf{N},$$
$$m < n \quad \text{si et seulement si} \quad mk < nk \quad \text{pour un} \quad k \in \mathbf{N}, k \neq 0.$$

On peut également remplacer le «pour un» par «pour tout» dans ces deux propositions. Les démonstrations sont laissées au lecteur. Il est également facile de démontrer que si m et n sont des entiers naturels et si $m \leqq n \leqq m + 1$, alors $m = n$ ou $n = m + 1$. Nous en laissons la démonstration au lecteur.

Démontrons maintenant la première propriété mentionnée au chapitre I, §2 concernant les entiers, à savoir le fait que $\mathbf{N}$ est bien ordonné: *tout sous-ensemble non vide F de $\mathbf{N}$ possède un plus petit élément.*

Pour le voir, considérons T la partie de $\mathbf{N}$ constituée de tous les n tels que $n \leqq x$ pour tout $x \in E$. L'élément $0 \in T$ et $T \neq \mathbf{N}$. Il existe par conséquent $m \in T$ tel que $m + 1 \notin T$ (par récurrence!). Alors $m \in E$ (sinon $m < x$ pour tout $x \in E$, ce qui est impossible). Il est donc clair que m est le plus petit élément de E, comme on le voulait.

Nous avons supposé connues au chapitre VIII les propriétés des cardinaux finis. Nous allons maintenant les démontrer. Soit, pour tout entier naturel $n \neq 0$, J_n l'ensemble des entiers naturels x tels que $1 \leqq x \leqq n$.

Si $n = 1$, alors $J_n = \{1\}$, et il n'y a qu'une seule application de J_1 dans lui-même. Cette application est évidemment bijective. Rappelons qu'on dit que deux ensembles A et B ont même *cardinal* s'il existe une bijection de A dans B. Puisqu'un produit de composition de bijections est une bijection, il s'ensuit que si

$$\operatorname{card}(A) = \operatorname{card}(B) \qquad \text{et} \qquad \operatorname{card}(B = \operatorname{card}(C),$$

alors $\operatorname{card}(A) = \operatorname{card}(C)$.

Soit m un entier naturel $\geqq 1$ et soit $k \in J_{m'}$. Il existe dans ces conditions une bijection entre

$$J_{m'} - \{k\} \qquad \text{et} \qquad J_m$$

on définit $f : J_{m'} - \{k\} \to J_m$, de façon évidente par

$$\begin{aligned} f &: x \mapsto x & \text{si} \quad & x < k, \\ f &: x \mapsto \sigma^{-1}(x) & \text{si} \quad & x \geqq k. \end{aligned}$$

On définit $g : J_m \to J_{m'} - \{k\}$, par

$$\begin{aligned} g &: x \mapsto x & \text{si} \quad & x < k, \\ g &: x \mapsto \sigma(x) & \text{si} \quad & x \geqq k. \end{aligned}$$

Alors $f \circ g$ et $g \circ f$ sont les identités correspondantes, de telle sorte que f et g sont des bijections.

Nous en déduisons que, *pour tous les entiers naturels $m \geqq 1$, si*

$$h : J_m \to J_m$$

est une injection, alors h est une bijection. En effet, cela est vrai pour $m = 1$, et par hypothèse de récurrence, supposons que l'assertion soit vraie pour un $m \geqq 1$. Soit

$$\varphi : J_{m'} \to J_m$$

une injection. Soit $r \in J_{m'}$ et soit $s = \varphi(r)$. On peut alors définir une application

$$\varphi_0 : J_{m'} - \{r\} \to J_{m'} - \{s\}$$

par $x \mapsto \varphi(x)$. Le cardinal de chaque ensemble $J_{m'} - \{r\}$ et $J_{m'} - \{s\}$ est le même que celui de J_m. Par hypothèse de récurrence, φ_0 est une bijection, et par suite w aussi, comme on le voulait.

Nous déduisons de ce qui précède que si $1 \leqq m < n$, *une application*

$$f : J_n \to J_m$$

ne peut pas être injective. Sinon, d'après ce qu'on a vu, $f(J_m) = J_m$ et $f(n) = f(x)$ pour un x tel que $1 \leqq x \leqq m$, de sorte que f n'est pas injective.

On dit, étant donné un ensemble A, que $\text{card}(A) = n$ (ou le cardinal de A est n, ou que A possède n éléments) pour un entier naturel $n \geqq 1$, s'il existe une bijection de A avec J_n. D'après les résultats précédents, un tel entier n est uniquement déterminé par A. On dit aussi que A est de cardinal 0 s'il est vide. On dit que A est *fini* s'il est de cardinal n pour un entier naturel n. C'est alors un simple exercice que de montrer les propositions suivantes:

Si A et B sont des ensembles finis et si $A \cap B$ est vide, alors

$$\text{card}(A) + \text{card}(B) = \text{card}(A \cup B).$$

De plus,

$$\text{card}(A)\,\text{card}(B) = \text{card}(A \times B).$$

Nous en laissons la démonstration au lecteur.

§2. *Les entiers*

Ayant les entiers naturels, nous voulons définir les *entiers* (ou *entiers relatifs*, ou encore *entiers rationnels*, N.d.T.). Nous allons le faire comme on le fait à l'école primaire: il n'y a pas de meilleure façon.

Pour tout entier naturel $n \neq 0$, nous choisissons un nouveau symbole noté $-n$, et désignons par $\mathbf{Z}$ l'ensemble constitué de la réunion de $\mathbf{N}$ et de tous les symboles $-n$ pour $n \in \mathbf{N}, n \neq 0$. Il nous faut définir l'addition sur $\mathbf{Z}$. Si $x, y \in \mathbf{N}$ l'addition est la même que celle de $\mathbf{N}$. On pose pour tout $x \in \mathbf{Z}$,

$$0 + x = x + 0 = x.$$

Cela est compatible avec l'addition définie au §1 pour $x \in \mathbf{N}$.

Soient $m, n \in \mathbf{N}$, tous deux non nuls. Si $m = n + k$ pour un $k \in \mathbf{N}$, on pose

(a) $m + (-n) = (-n) + m = k$;
(b) $(-m) + n = n + (-m) = -k$ si $k \neq 0$, et $= 0$ si $k = 0$;
(c) $(-m) + (-n) = -(m + n)$.

Etant donnés $x, y \in \mathbf{Z}$, non tous deux dans $\mathbf{N}$, une au moins des situations (a), (b) ou (c) s'applique à leur somme.

Il est alors fastidieux mais facile de vérifier que $\mathbf{Z}$ est un groupe additif.

Définissons maintenant la multiplication dans $\mathbf{Z}$. Si $x, y \in \mathbf{N}$, on utilise la multiplication de $\mathbf{N}$. Pour tout $x \in \mathbf{Z}$, on pose $0x = x0 = 0$.

Soient $m, n \in \mathbf{N}$ et tous deux non nuls. On pose:

$$(-m)n = n(-m) = -(mn) \quad \text{et} \quad (-m)(-n) = mn.$$

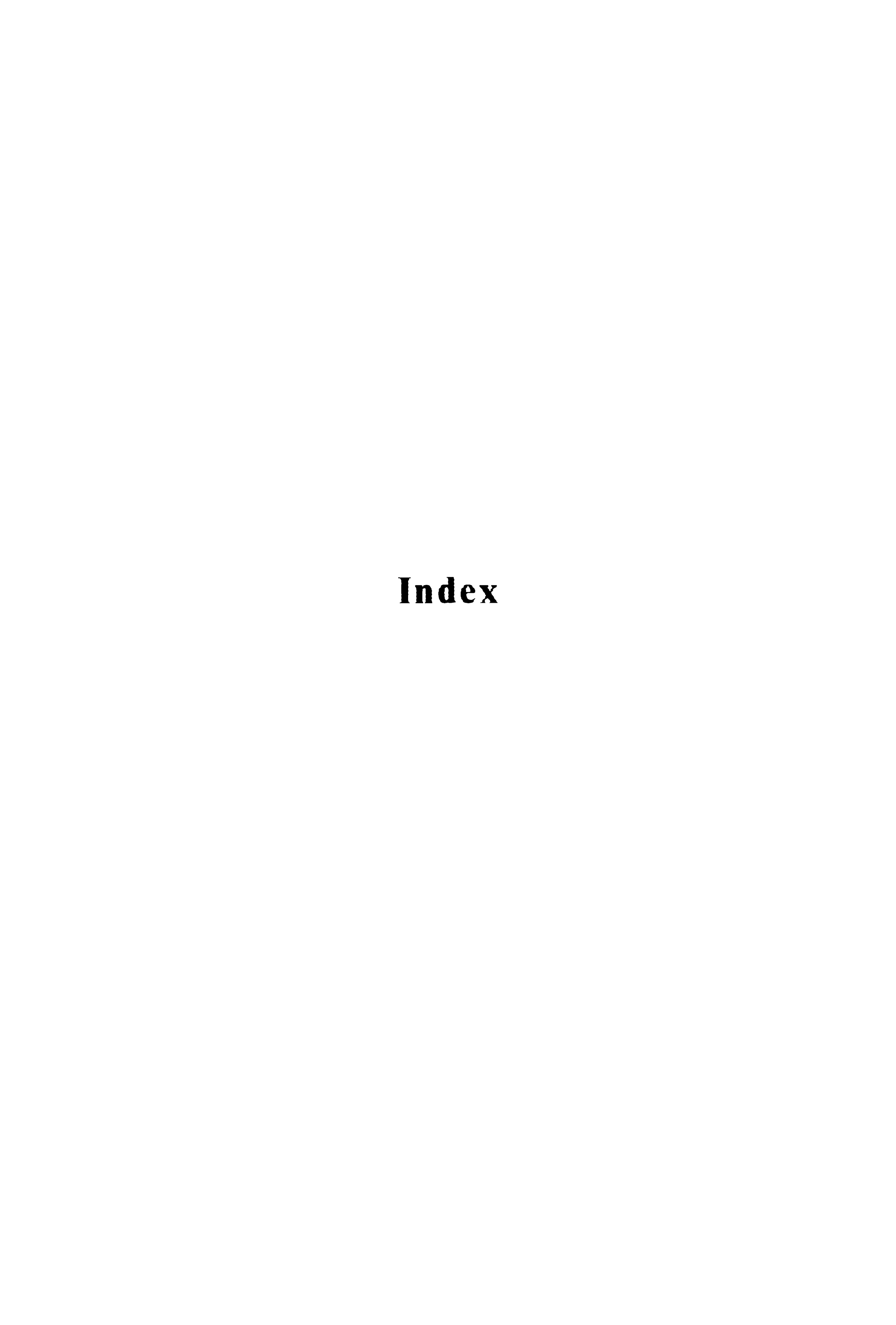
C'est alors une démonstration de routine qui prouve que $\mathbf{Z}$ est un anneau commutatif, en fait intègre dont l'élément unité est l'élément 1 de $\mathbf{N}$. On obtient de cette façon les entiers.

Remarquons que $\mathbf{Z}$ est un anneau ordonné au sens du chapitre VIII, §1 parce que l'ensemble des entiers naturels $\neq 0$ vérifie toutes les conditions données dans ce chapitre, comme on le constate directement à partir de nos définitions de la multiplication et de l'addition.

§3. *Ensembles infinis*

Un ensemble A est dit *infini* s'il n'est pas fini (et en particulier, pas vide).

Nous allons démontrer que *tout ensemble infini contient un ensemble dénombrable.* Choisissons un élément x dans chaque partie T non vide de A. Nous allons démontrer par récurrence, que, pour tout entier positif n, on peut trouver des éléments $x_1, \ldots, x_n \in A$, déterminés de façon unique, tels que $x_1 = x_A$ si l'élément choisi correspond à l'ensemble A lui-même, et tels que pour tout $k = 1, \ldots, n - 1$, l'élément x_{k+1} est l'élément choisi dans le complementaire de $\{x_1, \ldots, x_k\}$. Quand $n = 1$, il n'y a rien à démontrer. Supposons l'assertion démontrée pour $n > 1$. Soit alors x_{n+1} l'élément choisi dans le complémentaire de $\{x_1, \ldots, x_n\}$. Si $x_1, \ldots, x_n$ sont déterminés de façon unique, x_{n+1} aussi. Cela prouve ce que nous voulons. Puisque les éléments $x_1, \ldots, x_n$ sont distincts, pour tout n, il s'ensuit, en particulier, que la partie de A constituée de tous les x_n est un ensemble dénombrable, tel qu'on en cherchait un.



Index

Index

Les locutions contenant un adjectif sont classées à celui-ci

DEPOT LEGAL PREMIER TRIMESTRE 1976
INTEREDITIONS, PARIS S. A.
IMPRIME AUX ETATS-UNIS